INHERITANCE OF VOLUME AND RUST RESISTANCE IN SLASH PINE

By

MARK JOHN DIETERS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

ACKNOWLEDGEMENTS

I would like to express my gratitude to Drs. T.L. White, M.A. DeLorenzo, G.R. Hodge, R.C. Littell, and D.L. Rockwood for the time and effort they expended serving as members of my advisory committee. Particularly, I wish to thank Dr. T.L. White for his concern and guidance throughout the three and a half years during which he acted as chairman of my advisory committee.

In addition I would like to thank Dr. T.L. White and the Cooperative Forest Genetics Research Program (CFGRP) for providing financial support, members of the CFGRP who provided data from there tests, and Mr. G. Powell for his many hours assistance in cleaning and checking data.

I would also like to acknowledge Drs. M.U. Slee and D.G. Nikles for their encouragement of my career in forest tree improvement and their help when I was considering whether or not to pursue graduate education in quantitative genetics.

Finally, I want to thank my wife for her understanding and encouragement over these last three and a half years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

INHERITANCE OF VOLUME AND RUST RESISTANCE IN SLASH PINE

By

Mark John Dieters

August 1994

Chairperson: Timothy L. White
Major Department: School of Forest Resources and Conservation

Data from 171 full-sib tests of slash pine (*Pinus elliottii* Engelm. var *elliottii*) were used to obtain REML (restricted maximum likelihood) estimates of variance components and genetic parameters for volume and rust resistance. These tests include progeny of over 700 first-generation parents, more than 2100 families, and approximately 170,000 trees, measured at multiple ages between four and fifteen years.

The theory, derivation and application of two approximate methods for estimating the variance of heritability are described. Using an independent subset of the data, both the Dickerson and Taylor series approximations gave similar estimates of the variance and were of the same magnitude as an empirical estimate of the variance, regardless of the method used to estimate variance components (ANOVA-based or REML), test size, or the trait (continuous or binomial). The Dickerson approximation was, however, slightly more conservative.

The unbiased (paired site) heritability of volume was low: 0.07 at 5 years increasing to around 0.12 at 11 and 14 years, while the ratio of additive to dominance variance (σ_A^2/σ_D^2) was approximately 1.6 at 5 years, increasing to around 2.0 at 11 or more years. Genotype-by

environment interaction (GXE) was evident in volume, but relatively unimportant (type B genetic correlations increased from 0.6 at 5 years to 0.8 at 14 years).

The biased heritability of fusiform rust resistance was modelled as a quadratic function of mean rust incidence ($R^2=0.34$), maximizing at $h_b^2 = 0.195$, and 71.5% rust infection. At rust infection levels less than 66%, a simple linear function of the mean rust incidence ($R^2=0.43$), yielded unbiased heritability predictions which were consistent with the predicted h_b^2. Predicted type B genetic correlations ($R^2=0.18$) indicated that GXE is of little practical importance. The proportion of dominance variance in rust resistance was not related to rust incidence, age or any test parameter, and averaged 0.088 and 0.053 in single-site and paired-site analyses, respectively. Dominance was of little practical significance. Test age was not a significant factor affecting any genetic parameter of rust resistance, and all age-age genetic correlations are near 1.0.

Parameter estimates are discussed in terms of the current slash pine breeding and testing strategies. Changes to capture genetic gain from dominance variance in volume must be weighed against the additional costs.

CHAPTER 1
INTRODUCTION

Accurate and precise estimates of genetic parameters are very important to the continued efficient improvement of forest trees. These parameters are used at many stages in the breeding cycle (White 1987) for the development of breeding and testing strategies, for the prediction of breeding values, and for the prediction of genetic gains. Typically in forest trees, generation intervals are relatively long (over ten years), so it is more important than in most other crops, to have reliable genetic parameter estimates to ensure optimal genetic gains per unit time.

Although the importance of genetic parameters is widely recognized, as indicated by the large number of published parameter estimates in many tree species, estimates often come from a small number of genetic tests (often one test) using sub-optimal estimation techniques. Genetic parameters estimated from single tests are known to be upwardly biased due to the presence of confounding genotype-by-environment interaction (Comstock and Moll 1963), and the difficulty of estimating genetic parameters means that a large amount of data is required to obtain useful estimates. In addition, forest genetics tests often suffer from varying levels of imbalance, making the use of traditional Analysis of Variance (ANOVA) based methods of estimating variance components less than optimal (Searle *et al.* 1992, Huber 1993). In simulation studies Restricted Maximum Likelihood (REML) estimation has been shown to be generally more appropriate than ANOVA-based estimation methods when data are unbalanced, under a range of types and levels of imbalance commonly found in forest genetic tests (Huber 1993).

This dissertation utilizes individual tree data collected from 171 full-sib tests of slash pine (*Pinus elliottii* Engelm. var *elliottii*) measured for volume and the incidence of fusiform rust at multiple ages between 4 and 15 years to estimate variance components and genetic parameters. These tests were established, maintained and measured by members of the Cooperative Forest Genetic Research Program (CFGRP), a cooperative tree improvement program based at the University of Florida and composed of both industrial and government organizations in the southeastern United States. This very large data set (involving over 700 first-generation parents, represented by over 2100 full-sib families, and approximately 170,000 trees) offers unique opportunities for the estimation of genetic parameters in slash pine. Few other forest tree improvement programs have data sets of this magnitude available for parameter estimation.

Heritability (the ratio of additive genetic variance to total phenotypic variance) is one of the most important genetic parameters and is often used to summarize information from genetic tests. Heritability estimates are used in strategy development, in the prediction of genetic gains, in the formulation of selection indices and in the prediction of breeding values (Falconer 1989, White and Hodge 1989). Given the importance of heritability estimates, it is natural to desire an estimate of the variance of the estimated heritability. To achieve this, an approximate method based on a Taylor series is frequently used to estimate the variance of the heritability (Kempthorne 1957, Becker 1975, Hallauer and Miranda 1988). The extensive nature of the data available to this study made it possible to investigate the reliability of the Taylor series approximation compared to a simpler approximation (Dickerson 1969) and an empirical estimate of the variance for a continuous trait (tree volume) and a Bernoulli trait (rust resistance). Therefore, as the first part of this dissertation the theory, derivation, and application of these approximate methods are fully described and the utility of these approximations was tested using

a subset of the data analyzed using both ANOVA-based and REML estimation methods (Chapter 2).

Genetic parameters (heritability, proportion of dominance, type B genetic correlations, and age-age genetic correlations) for both tree volume and rust resistance were estimated using GAREML (Huber 1993) to provide REML (Patterson and Thompson 1971) estimates from both single and paired test analyses (Chapters 3 and 4). No previous study in the forest genetics literature has utilized a data set comparable in magnitude to that from which the estimates in this dissertation were derived. This study provides reliable estimates of the importance of dominance variance in tree volume and rust resistance of slash pine for the first time, and it is the most comprehensive published study of dominance variance in any of the southern pines.

Average parameter estimates for volume in slash pines at 5, 8, 11, and 14 years are presented, and the influence of site factors on these parameters is investigated. For resistance to fusiform rust in slash pine it was possible to develop models to predict heritability and type B genetic correlations based on simple functions of the mean level of rust observed in the tests. Previously there were some indications that parameter estimates of rust resistance were influenced by the level of rust infection (Rockwood and Goddard 1973, Sohn and Goddard 1979, White and Hodge 1989); however, all such studies in slash pine have been based on a much smaller number of tests. Implications of these parameter estimates for the continued genetic improvement of slash pine by the CFGRP are discussed.

CHAPTER 2
DERIVATION AND APPLICATION OF APPROXIMATE VARIANCES OF VARIANCE COMPONENTS AND THEIR RATIOS IN GENETIC TESTS

Introduction

The results of quantitative genetics tests are commonly summarized in terms of estimated variance components and ratios of these variance components such as heritability and genetic correlation. When examining such statistics, it is natural to inquire about their precision; hence, standard errors are customarily reported for the variance component estimates and ratios of interest. The estimation of such standard errors is not simple, and in the case of the variance of a ratio of variance components, no closed form expression exists.

Knapp *et al.* (1985) present exact confidence intervals for family mean heritability from balanced data, and nonparametric confidence intervals can also be estimated (Knapp *et al.* 1989). However, most commonly formulae based on a Taylor series approximation are used to estimate standard errors of ratios of variance components (Tallis 1959, Swiger *et al.* 1964, Kempthorne 1957, Becker 1975, Namkoong 1979, Bulmer 1980, Baker 1986, Hallauer and Miranda 1988, Falconer 1989). Although the theory of this Taylor series approximation is accessible in advanced calculus and statistics texts (Kaplan 1952; Kendall *et al.* 1987; Searle *et al.* 1992), neither the derivation nor the underlying assumptions required in the derivation and application of these formulae have been presented in either the forestry or the genetics literature. Failure to clearly

understand the parameters and the assumptions implicit in these formulae can lead to erroneous and inappropriate application.

The objectives of this paper are twofold. First, to lay out the theory and assumptions needed to derive approximate standard errors of ANOVA-based and REML-based estimators of variance components and their ratios, for both balanced and unbalanced data. The second objective is to compare the approximate variances of ratios of variance components to empirically determined variances among from many separate estimates of these same ratios using real data. The ratios of variance components which will be used in this study are, i) single-site biased heritability (h_b^2) estimated using ANOVA and REML, and ii) unbiased heritability (h^2) from paired sites estimated using REML. The case studies will examine two traits from the Cooperative Forest Genetic Research Program (CFGRP) slash pine (*Pinus elliottii* Engelm. var *elliottii*) breeding program: resistance to fusiform rust (caused by the fungus *Cronartium quercuum* f. sp. *fusiforme*) which is a bernoulli trait, assessed as infected or rust-free, with moderate heritability, and tree volume, a lower heritability trait.

Theory

Variance of Variance Components — ANOVA Estimates

All ANOVA-based estimators of variance components rely on equating observed mean squares to their expectations and solving for the variance components. Consequently, ANOVA-based variance component estimates are linear combinations of the mean squares. When data are balanced, ANOVA-based estimates of variance components are unique, unbiased, and have the minimum variance among all unbiased quadratic estimators (Swallow and Monahan 1984; Khuri and Sahai 1985; Searle *et al.* 1992). These optimal statistical properties usually make ANOVA-based estimators the best choice when data are balanced. Although ANOVA-based variance

component estimates are derived from sums of squares in the ANOVA table, no assumption of normality is required to derive ANOVA estimators (Searle *et al.* 1992). However, in order to derive the usual estimator for the variance of variance component estimates, the traditional normality assumptions for analysis of variance are required. These normality assumptions are that all random effects in the model are normally distributed with a mean of zero, some finite variance, and zero covariances between effects (Searle *et al.* 1992, Winer *et al.* 1991).

In order to apply formulae for the variance of ratios of variance components, it is necessary to know the variance and covariance among the variance component estimates. To obtain this information we must first estimate the variance of the mean squares in the ANOVA. Once the variance of the mean squares is known, it is relatively simple to estimate the variances and covariances of the variance component estimates.

Variance of a mean square

Searle *et al.* (1992) derive the variance of a mean square in Appendix S.3 of their book, however a few steps are omitted. A simplification of their derivation is therefore given here for completeness. Let SS be a sum of squares with f degrees of freedom and corresponding mean square, MS=SS/f with expected values E(SS)=fE(MS). Under the assumption of normality with balanced data, the sum of squares divided by the expected value of its mean square is distributed as a central chi-square with f degrees of freedom.

The variance of a random variable with a central chi-squared distribution is two times its degrees of freedom; hence

$$Var\left[\frac{SS}{E(MS)}\right] = 2f$$

and since the E(MS) is a constant, this is equal to

$$\frac{Var(SS)}{[E(MS)]^2} = 2f$$

and thus,

$$Var(SS) = 2f[E(MS)]^2.$$

Further, because SS = fMS

$$Var(SS) = Var(fMS) = f^2 Var(MS),$$

therefore,

$$Var(MS) = \frac{Var(SS)}{f^2} = \frac{2f[E(MS)]^2}{f^2} = \frac{2[E(MS)]^2}{f} \qquad (2\text{-}1)$$

Hence we have shown that the variance of a mean square is a function of the expected value of that mean square and its degrees of freedom. However, the expected value of the mean square is unknown, and so we must seek to re-express equation 2-1 in terms of the observed value of the mean square.

By definition, the variance of any random variable, say X, is

$$Var(X) = E[X - E(X)]^2 = E(X^2) - [E(X)]^2$$

hence,

$$Var(MS) = E(MS^2) - [E(MS)]^2. \qquad (2\text{-}2)$$

By equating these two expressions (equations 2-1 and 2-2) for the variance of a mean square, we have the following,

$$\frac{2[E(MS)]^2}{f} = E(MS^2) - [E(MS)]^2$$

and rearranging,

$$E(MS^2) = [E(MS)]^2 + \frac{2[E(MS)]^2}{f}$$

Multiplying through by f and collecting like terms yields

$$fE(MS^2) = (f + 2)[E(MS)]^2$$

and so,

$$\frac{E(MS^2)}{(f+2)} = \frac{[E(MS)]^2}{f}$$

which, because $(f + 2)$ is a constant, is numerically equivalent to

$$E\left[\frac{(MS^2)}{(f+2)}\right] = \frac{[E(MS)]^2}{f}$$

Therefore $(MS^2)/(f + 2)$ is an unbiased estimator of $[E(MS)]^2/f$. If we substitute this unbiased estimator into equation 2-1 in place of $[E(MS)]^2/f$, we find that

$$Var(MS) = \frac{2(MS^2)}{(f+2)} \qquad (2\text{-}3)$$

The expression in equation 2-3 for the variance of a mean square is the same as that given on page 246 of Kempthorne (1957).

Variance of a variance component

The variance of an ANOVA-based variance component estimate is derived by using the fact that all such variance component estimates are linear functions of the observed mean squares. The variance of any linear function of two random variables X and Y, such as

$$U = aX + bY$$

where a and b are constants, is

$$\mathrm{Var}(U) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X,Y).$$

This formula can be readily extended to the multivariate case. Let $X_1, X_2, \ldots, X_n$ be random variables, and $a_1, a_2, \ldots, a_n$ be constants. Then for $U = \Sigma_i^n a_i X_i$,

$$\mathrm{Var}(U) = \Sigma_i^n a_i^2\mathrm{Var}(X_i) + 2\Sigma\Sigma_{i<j} a_i a_j \mathrm{Cov}(X_i, X_j), \qquad (2\text{-}4)$$

where the double summation is over all pairs (i,j) with $i<j$ (Mendenhall *et al.* 1990).

Equations 2-3 and 2-4 can be used to estimate the variance of any ANOVA-based variance component estimate. For example let σ_1^2 be estimated by,

$$\hat{\sigma}_1^2 = (MS_1 - MS_2)/c,$$

where MS_1 and MS_2 are mean squares from the analysis of variance with degrees of freedom f_1 and f_2 respectively, and c is a constant. Then,

$$\mathrm{Var}(\hat{\sigma}_1^2) = \mathrm{Var}((MS_1 - MS_2)/c) = 1/c^2\ \mathrm{Var}(MS_1 - MS_2)$$

using equation 2-4,

$$\mathrm{Var}(\hat{\sigma}_1^2) = 1/c^2\ [\mathrm{Var}(MS_1) + \mathrm{Var}(MS_2) - 2\mathrm{Cov}(MS_1, MS_2)]$$

and applying equation 2-3,

$$\mathrm{Var}(\hat{\sigma}_1^2) = 2/c^2\ [(MS_1^2/(f_1 + 2)) + (MS_2^2/(f_2 + 2))],$$

if we assume that the covariance between the mean squares is zero. With balanced data the mean squares are orthogonal (Steel and Torrie 1981), which in turn means that there is no covariance between the mean squares (Namkoong 1979). Nevertheless, with the assumed variance-covariance structure, and the assumption of normality, all the mean squares are independent (Searle *et al.* 1992). These assumptions are usually reasonable in most designed experiments where random samples are drawn from normal populations and treatments are assigned randomly to experimental units (Winer *et al.* 1991).

This procedure can be generalized. The i^{th} variance component can be estimated as a linear combination of the l mean squares, each with f_i degrees of freedom,

$$\hat{\sigma}_i^2 = (a_1 MS_1 + a_2 MS_2 + \ldots + a_k MS_l),$$

or equivalently,

$$\hat{\sigma}_i^2 = \Sigma_i^l a_i MS_i,$$

where the a's are constants, equivalent to $1/c$ and $-1/c$ in the previous example. And the variance of this linear function of the mean squares is estimated using

$$\mathrm{Var}(\hat{\sigma}_i^2) = \mathrm{Var}(\Sigma_i^l a_i MS_i) = \Sigma_i^l \mathrm{Var}(a_i MS_i) = \Sigma_i^l a_i^2 \mathrm{Var}(MS_i).$$

By using equation 2-3 to estimate the variance of MS_i, we obtain

$$\hat{Var}(\hat{\sigma}^2) = \sum_{i=1}^{l} \left[\frac{2 a_i^2 (MS_i)^2}{f_i + 2} \right] \qquad (2\text{-}5)$$

which is the same as the formula used by Schaffer and Usanis (1969, p.26) to estimate the standard deviation of the estimated variance components.

Searle *et al.* (1992, pp. 128-138) present a useful matrix formulation for estimating variance components and their variances from an ANOVA table. If $\mathbf{P}$ is an lxl matrix of the coefficients of the variance components in the expectation of the mean squares, $\mathbf{m}$ an lx1 vector of mean squares from the analysis of variance, and $\boldsymbol{\sigma}^2$ an lx1 vector of the variance components to be estimated (l is the number of mean squares as above), then the ANOVA estimator of $\boldsymbol{\sigma}^2$, $\hat{\boldsymbol{\sigma}}^2$, is obtained as follows: by definition,

$$E(\mathbf{m}) = \mathbf{P}\boldsymbol{\sigma}^2, \text{ and}$$

$$\mathbf{m} = \mathbf{P}\hat{\boldsymbol{\sigma}}^2.$$

Therefore, provided that $\mathbf{P}$ is nonsingular (i.e., has an inverse),

$$\hat{\boldsymbol{\sigma}}^2 = \mathbf{P}^{-1}\mathbf{m} \qquad (2\text{-}6)$$

and the variance of $\hat{\sigma}^2$ is

$$\mathrm{Var}(\hat{\sigma}^2) = \mathrm{Var}(\mathbf{P}^{-1}\mathbf{m}) = \mathbf{P}^{-1}\mathrm{Var}(\mathbf{m})\mathbf{P}^{-1'}$$

And the variance of $\hat{\sigma}^2$ is estimated by

$$\hat{\mathrm{Var}}(\hat{\sigma}^2) = \mathbf{P}^{-1}\mathbf{D}\mathbf{P}^{-1'}, \qquad (2\text{-}7)$$

where $\mathbf{P}^{-1'}$ is the transpose of the inverse of $\mathbf{P}$, and $\mathbf{D}$ is a diagonal (lxl) matrix with the variances of the individual mean squares (i.e., $2(\mathrm{MS}_i)^2/(f_i+2)$, as previously discussed) on the diagonal and zeros elsewhere. The lxl matrix $\hat{\mathrm{Var}}(\hat{\sigma}^2)$ has the variances of the variance components on the diagonal and the covariances between the variance components on the off-diagonals. The matrix representation in equations 2-6 and 2-7 is especially useful because it can be programmed readily.

In summary, for balanced data it is possible to derive the variance of ANOVA-based variance component estimates, requiring only the traditional normality assumptions, plus the assumption of zero covariance between the mean squares. However in the case of unbalanced data using a random or mixed model, in general the mean squares do not have chi-squared distributions, nor any multiple of a chi-squared distribution (Searle *et al.* 1992). The validity of applying this methodology to variance components estimated from unbalanced data can not be readily determined theoretically but at best can be expected to provide approximate variances when the data are nearly balanced.

Variance of Variance Components — REML Estimates

For unbalanced data, REML can offer significant advantages over ANOVA-based estimators: REML estimates are unique and nonnegative and have some useful large-sample statistical properties, while ANOVA-based estimates only retain the property of being unbiased (Swallow and Monahan 1984, Khuri and Sahai 1985, Searle *et al.* 1992). Hartley and Rao (1967)

showed that with certain restrictions REML-based estimates are asymptotically normal, consistent and asymptotically efficient as the design size increases. However these properties have subsequently been verified under fairly unrestrictive assumptions (Harville 1977).

The derivation of REML estimators requires that an underlying probability distribution for all random effects in the model be assumed. Commonly a multivariate normal distribution is assumed, because this assumption leads to solutions that can be readily handled mathematically, even when the data are unbalanced (Searle *et al.* 1992). Nevertheless, Banks *et al.* (1984) and Westfall (1987) have demonstrated that REML-based estimators are robust to violations of this assumption.

When using any maximum likelihood procedure, the asymptotic variance and covariances of the variance component estimates can be estimated by taking the inverse of the information matrix (Searle *et al.* 1992). The information matrix is the negative of the expectation of the matrix of second derivatives of the likelihood function, which can be derived readily as part of the estimation process. However, these are asymptotic or large-sample properties, and so are only approximate when applied to variance component estimates derived from finite (small) samples. Nevertheless many computer programs, such as PROC VARCOMP of the SAS system (SAS Institute Inc. 1988) and GAREML (Huber 1993) which perform REML estimation, also provide the asymptotic variance-covariance matrix of the estimates. These variances and covariances are used to provide a measure of the precision of REML estimates even when sample size is small, because no other estimates are available.

The approximate nature of the asymptotic properties of REML estimates derived from finite data has been demonstrated in simulation studies. Although REML estimates are asymptotically unbiased, in practice they are biased (Swallow and Monahan 1984, Westfall 1987, Khattree and Gill 1988, Huber 1993).

Variance of a Ratio

Unlike linear functions of random variables, no closed form expression exists for estimating the variance of a ratio of two random variables. Heritabilities and genetic correlations are ratios of estimated variance components and so are ratios of random variables. The variance of such ratios can, however, be approximated by making simplifying assumptions. Two approximate methods will be described here. The first is a simple method proposed by Dickerson (1969) that may have considerable merit but which is rarely used. The second method is based on a Taylor series approximation and is the method most commonly used in quantitative genetics.

Dickerson's method

Dickerson (1969) suggested a simple method for obtaining estimates of the variance of a ratio. Dickerson's method estimates the variance of the ratio cx_1/x_2 (where x_1 and x_2 are random variables and c is a constant) as

$$Var\left(\frac{cx_1}{x_2}\right) = \frac{c^2\ Var(x_1)}{x_2^2} \qquad (2\text{-}8)$$

Equation 2-8 is derived by treating the denominator, x_2, as a constant rather than as a random variable. The assumption that the denominator is a constant appears to be justifiable in the case of heritability estimates. Narrow sense heritability is the ratio between the additive variance and the total phenotypic variance. Estimates of the total phenotypic variance are usually much more precise than are estimates of additive variance. Thus, for most practical purposes the estimate of phenotypic variance could be considered as a known parameter rather than an estimate of that parameter. Dickerson (1969) states that although this method is conservative (i.e., slightly overestimates the variance), it gives results that agree closely with those obtained from the Taylor series approximation.

Taylor series approximation to a function in 2-space

The Taylor series approximations has its origin in calculus and may be used to approximate many continuous functions (Kaplan 1952, p.357). In order to adequately understand the derivation of formulae often cited for the variance of a ratio, it is first necessary to examine how the Taylor series expansion is used in calculus. To this end we will examine the application of the Taylor series to a polynomial in 2-space.

An n^{th} degree polynomial in 2-space defined by the xy-plane is customarily written in the form

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots + a_nx^n,$$

that is, in terms of powers of x (Goodman 1969). For example, a linear function ($n = 1$) is represented by the first two terms of this series, as is well known. An alternative and equivalent representation is in terms of powers of $(x - a)$, giving the following form

$$f(x) = c_0 + c_1(x - a) + c_2(x - a)^2 + c_3(x - a)^3 + \ldots + c_n(x - a)^n, \qquad (2\text{-}9)$$

where a is some constant. Any polynomial has a unique expansion of this form, about any given point a (Goodman 1969). The Taylor series is equivalent to a polynomial of infinite degree centered at a, i.e., equation 2-9 with $n = \infty$, and the coefficients c_0, c_1, c_2, ... c_n, ... are determined by the value of $f(x)$ and its derivatives evaluated at $x = a$. For simplicity, we use the notation $f(a)$ for $f(x)$ evaluated at the point $x = a$, $f'(a)$ for the first derivative of $f(x)$ with respect to x, evaluated at the point $x = a$, etc. If we expand the first four terms of $f(x)$,

$$f(x) = c_0 + c_1(x - a) + c_2(x^2 - 2xa + a^2) + c_3(x^3 - 3x^2a + 3xa^2 + a^3) + \ldots$$

Hence, evaluating $f(x)$ at $x = a$, we find $f(a) = c_0$, and similarly for the first derivative evaluated at $x = a$,

$$f'(a) = \partial f(x)/\partial x \,|_{x=a} = c_1 + c_2(2x + 2a) + c_3(3x^2 - 6xa + 3a^2) + \ldots \,|_{x=a}$$

$$f'(a) = c_1 \quad \text{or} \quad c_1 = f'(a)/1!$$

and the second derivative evaluated at $x = a$,

$$f''(a) = \partial^2 f(x)/\partial x^2 \,|_{x=a} = 2c_2 + c_3(6x - 6a) + \ldots \,|_{x=a}$$

$$f''(a) = 2c_2 \quad \text{or} \quad c_2 = f''(a)/2!.$$

This can be generalized to the more common form of the Taylor series expansion:

$$f(x) = f(a) + f'(a)(x - a)/1! + f''(a)(x - a)^2/2! + \ldots + f^{(n)}(a)(x - a)^n/n! + \ldots$$

as found in calculus textbooks such as Kaplan (1952), Goodman (1969), and Murtha and Willard (1973). The Taylor series can also be expressed as a finite series (Kaplan 1952) with the addition of a remainder term,

$$f(x) = f(a) + f'(a)(x - a)/1! + f''(a)(x - a)^2/2! + \ldots$$
$$+ f^{(n)}(a)(x - a)^n/n! + R_n \qquad (2\text{-}10)$$

where R_n is the remainder term,

$$R_n = f^{(n+1)}(x_1)(x - a)^{n+1}/(n + 1)!$$

for some x_1 such that $a < x_1 < x$, or (if $x < a$) $x < x_1 < a$.

The application of the Taylor series expansion (equation 2-10) to the approximation of a polynomial can be better understood by reference to a simple second degree polynomial (n=2) such as, $f(x) = x^2$ (Figure 2-1). Let a be some point close to x, with x_1 intermediate between a and x, then

$f(a)$ = the value of $f(x)$ when $x = a$, i.e., a^2

$f'(a)$ = the slope of the line, tangent to the point $f(a)$, i.e., $2a$

$f''(x_1) = 2$, and

$(x - a)$ is taken to be a small number, Δ.

Point A on Figure 2-1, equal to $a^2 + 2a\Delta$, represents a first order approximation, and point B which is equal to $a^2 + 2a\Delta + \Delta^2$ is the second order approximation. Note that because

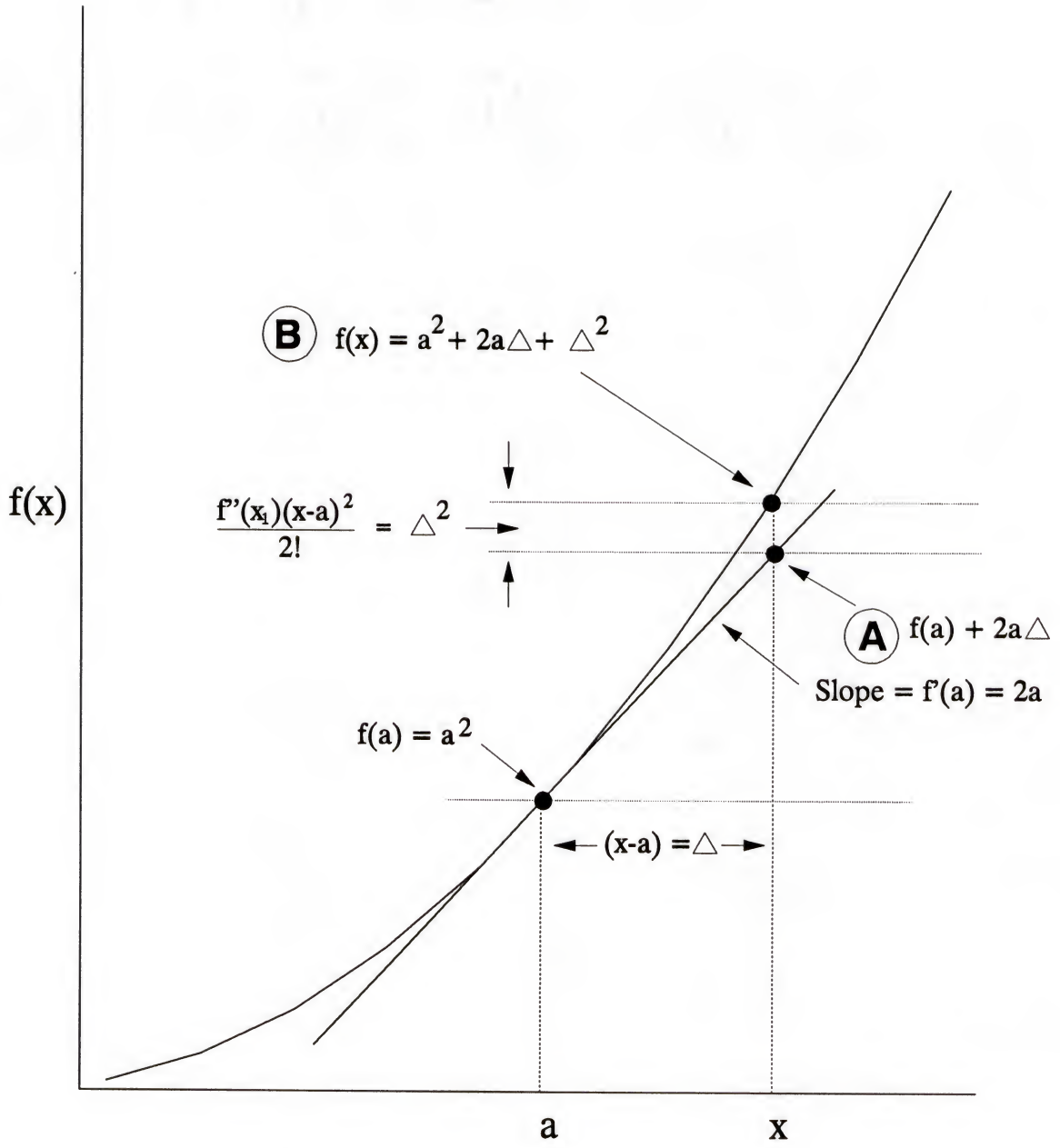

Figure 2-1. Plot of the second degree polynomial, $f(x) = x^2$. Points A and B are those estimated by the first- and second-order approximations of the Taylor series expansion. Points a and x can be any points on the x-line, and x_1 is chosen such that it lies between a and x.

this is only a second degree polynomial, it can be represented exactly in terms of the first three terms of the Taylor series expansion (a second order approximation), thus

$$f(x) = a^2 + 2a\Delta + (2\Delta^2)/2! = a^2 + 2a\Delta + \Delta^2.$$

Now consider the multivariate case and the function $g(x_1, x_2, \ldots x_k)$. If $g_i'(\theta)$ is $\partial g(x_1, x_2, \ldots x_k)/\partial x_i$ evaluated at $\theta_1, \theta_2, \ldots \theta_k$, thus from Kendall *et al.* (1987), the Taylor series is

$$g(x_1, x_2, \ldots x_k) = g(\theta) + \Sigma_i\, g_i'(\theta)(x_i - \theta_i) + O(n^{-r}) \qquad (2\text{-}11)$$

where θ has been substituted for a in equation 2-10,

$g(\theta)$ is equivalent to $f(a)$ in equation 2-10,

$\Sigma_i\, g_i'(\theta)(x_i - \theta_i)$ is equivalent to $f'(a)(x - a)/1!$ in equation 2-10, and

$O(n^{-r})$ is equivalent to the remainder term, R_n, in equation 2-10.

The formula presented in equation 2-11 is the first order expansion of the Taylor series. If all the first derivatives are zero, then further terms in the Taylor expansion must be used (Kendall *et al.* 1987). All further discussion of the Taylor series will be restricted to the first order expansion, because this (equation 2-11) forms the base most commonly used for the derivation of the variance of a ratio of two random variables.

Taylor series approximation to a ratio of two random variables

Standard quantitative genetics texts such as Kempthorne (1957, p.246), Becker (1975, p.45), and Namkoong (1979, p.232) present without proof or derivation formulae for the variance of the ratio of two random variables based on the Taylor series expansion which are all numerically equivalent to

$$\mathrm{Var}(g) \simeq \mathrm{Var}(x_1)/\theta_2^2 + \theta_1^2\mathrm{Var}(x_2)/\theta_2^4 - 2\theta_1\mathrm{Cov}(x_1,x_2)/\theta_2^3, \qquad (2\text{-}12)$$

where $g = x_1/x_2$, $\mathrm{E}(x_1) = \theta_1$, and $\mathrm{E}(x_2) = \theta_2$.

In the following discussion equation 2-12 is derived from the first order expansion of the Taylor series given in equation 2-11, and for brevity $g(x_1, x_2, \ldots, x_k)$ will be referred to simply

as $g(x)$. In order to do this it is necessary to derive i) the expected value of $g(x)$ and ii) the variance of $g(x)$, which are in turn used to derive the formula for the variance of a ratio given in equation 2-12.

The expected value of $g(x)$ is derived as follows:

$$\begin{aligned}
\mathrm{E}[g(x)] &= \mathrm{E}[g(\theta) + \Sigma_i\, g_i'(\theta)(x_i - \theta_i) + O(n^{-r})] \\
&= \mathrm{E}[g(\theta)] + \mathrm{E}[\Sigma_i\, g_i'(\theta)(x_i - \theta_i)] + \mathrm{E}[O(n^{-r})] && \text{Step 1} \\
&= g(\theta) + \Sigma_i\, \mathrm{E}[g_i'(\theta)(x_i - \theta_i)] + O(n^{-r}) && \text{Step 2} \\
&= g(\theta) + \Sigma_i\, g_i'(\theta)\mathrm{E}[(x_i - \theta_i)] + O(n^{-r}) && \text{Step 3} \\
&= g(\theta) + \Sigma_i\, g_i'(\theta)[\mathrm{E}(x_i) - \theta_i] + O(n^{-r}) && \text{Step 4} \\
&= g(\theta) + O(n^{-r}) && (2\text{-}13)
\end{aligned}$$

Initially, equation 2-11 is substituted for $g(x)$, and Step 1 proceeds by distributing the expectation operator with respect to addition. In Step 2, $g(\theta)$ and $O(n^{-r})$ are constants and so are equal to their own expectations, and in a similar manner to Step 1 the expectation operator can be placed inside the summation over i. Steps 3 and 4 use the fact that $g_i'(\theta)$ and θ_i are constants. And in the final step, the expression $\mathrm{E}(x_i) - \theta_i = 0$, because $\mathrm{E}(x_i) = \theta_i$.

The variance of $g(x)$ is derived as follows:

$$\begin{aligned}
\mathrm{Var}[g(x)] &= \mathrm{Var}[g(\theta) + \Sigma_i\, g_i'(\theta)(x_i - \theta_i) + O(n^{-r})] \\
&= \mathrm{Var}[g(\theta)] + \mathrm{Var}[\Sigma_i\, g_i'(\theta)(x_i - \theta_i)] + \mathrm{Var}[O(n^{-r})] + \text{Covariances} && \text{Step 1} \\
&= \mathrm{Var}[\Sigma_i\, g_i'(\theta)(x_i - \theta_i)] + O(n^{-r})^* && \text{Step 2} \\
&= \mathrm{Var}[g_1'(\theta)(x_1 - \theta_1) + g_2'(\theta)(x_2 - \theta_2) + \ldots + g_i'(\theta)(x_i - \theta_i) + \\
&\qquad \ldots + g_k'(\theta)(x_k - \theta_k)] + O(n^{-r})^* && \text{Step 3} \\
&= \mathrm{Var}[g_1'(\theta)(x_1 - \theta_1)] + \mathrm{Var}[g_2'(\theta)(x_2 - \theta_2)] + \ldots + \mathrm{Var}[g_i'(\theta)(x_i - \theta_i)] + \\
&\qquad \ldots + \mathrm{Var}[g_k'(\theta)(x_k - \theta_k)] + 2\mathrm{Cov}[g_1'(\theta)(x_1 - \theta_1),\, g_2'(\theta)(x_2 - \theta_2)] + \\
&\qquad 2\mathrm{Cov}[g_1'(\theta)(x_1 - \theta_1),\, g_3'(\theta)(x_3 - \theta_3)] + \ldots +
\end{aligned}$$

$$2\text{Cov}[g_1'(\theta)(x_1 - \theta_1),\ g_i'(\theta)(x_i - \theta_i)] + \ldots + O(n^{-r})^* \qquad \text{Step 4}$$

$$= \Sigma_{i=1}^{t} \text{Var}[g_i'(\theta)(x_i - \theta_i)] + \Sigma\Sigma_{i \neq j=1}^{t} \text{Cov}[g_i'(\theta)(x_i - \theta_i),\ g_j'(\theta)(x_j - \theta_j)] +$$

$$+ O(n^{-r})^* \qquad \text{Step 5}$$

$$= \Sigma_{i=1}^{t} \{g_i'(\theta)\}^2 \text{Var}[(x_i - \theta_i)] + \Sigma\Sigma_{i \neq j=1}^{t} g_i'(\theta) g_j'(\theta) \text{Cov}[(x_i - \theta_i),\ (x_j - \theta_j)]$$

$$+ O(n^{-r})^* \qquad \text{Step 6}$$

$$= \Sigma_{i=1}^{t} \{g_i'(\theta)\}^2 \text{Var}(x_i) + \Sigma\Sigma_{i \neq j=1}^{t} g_i'(\theta) g_j'(\theta) \text{Cov}(x_i,\ x_j)$$

$$+ O(n^{-r})^* \qquad (2\text{-}14)$$

The initial step in this derivation involves substitution of equation 2-11 for $g(x)$, and Step 1 proceeds by distributing the variance operator. Step 2 is based on the fact that $g(\theta)$ is a constant, and so has zero variance. In Step 2 the variance of $O(n^{-r})$ as well as the covariances between the x_i's and $O(n^{-r})$ are now represented by the single term, $O(n^{-r})^*$. The third step simply involves an expansion of the summation operator. Step 4 utilizes the theorem presented in equation 2-4. Step 5 is an alternative expression for Step 4 using summation notation. In Step 6 $g_i'(\theta)$ and $g_j'(\theta)$ are constants and so they can be taken outside the variance and covariance operators. The final step involves that fact that the θ_i's are constants and so contribute no variance or covariance.

It is now possible to derive a formula for the variance of a ratio: let $g(x_1, x_2) = x_1/x_2$. Thus, from equation 2-11, $g(x_1, x_2)$ is equal to

$$g(x_1, x_2) = g(x_1, x_2)|_{\theta 1 \theta 2} + \partial g(x_1, x_2)/\partial x_1|_{\theta 1 \theta 2} (x_1 - \theta_1) +$$

$$\partial g(x_1, x_2)/\partial x_2|_{\theta 1 \theta 2} (x_2 - \theta_2) + O(n^{-r})$$

where

$$g(x_1, x_2)|_{\theta 1 \theta 2} = \theta_1/\theta_2$$

$$\partial g(x_1, x_2)/\partial x_1|_{\theta 1 \theta 2} = 1/x_2|_{\theta 1 \theta 2} = 1/\theta_2, \text{ and}$$

$$\partial g(x_1, x_2)/\partial x_2|_{\theta 1 \theta 2} = -x_1/(x_2)^2|_{\theta 1 \theta 2} = -\theta_1/(\theta_2)^2.$$

Therefore,

$$g(x_1, x_2) = \theta_1/\theta_2 + (1/\theta_2)(x_1 - \theta_1) + (-\theta_1/(\theta_2)^2)(x_2 - \theta_2) + O(n^{-r}) \quad (2\text{-}15)$$

From equations 2-14 and 2-15 it is possible to estimate the variance of x_1/x_2 in the following manner

$$\begin{aligned}\mathrm{Var}(x_1/x_2) &= \Sigma_{i=1}^{2} \{g_i'(\theta)\}^2 \mathrm{Var}(x_i) + \Sigma\Sigma_{i\neq j=1}^{2} g_i'(\theta)g_j'(\theta) \mathrm{Cov}(x_i, x_j) + O(n^{-r}) \\ &= (1/\theta_2)^2 \mathrm{Var}(x_1) + (-\theta_1/(\theta_2)^2)^2 \mathrm{Var}(x_2) + \\ &\qquad 2(1/\theta_2)(-\theta_1/(\theta_2)^2) \mathrm{Cov}(x_1, x_2) + O(n^{-r}) \\ &= \mathrm{Var}(x_1)/\theta_2^2 + \theta_1^2\mathrm{Var}(x_2)/\theta_2^4 - 2\theta_1\mathrm{Cov}(x_1, x_2)/\theta_2^3 + O(n^{-r}) \quad (2\text{-}16)\end{aligned}$$

If we assume that $O(n^{-r})$ is negligible, then equation 2-16 is the same as equation 2-12. Thus the three assumptions required are i) all second and higher order terms contained in $O(n^{-r})$ are negligible, ii) all the covariances between the x_i's and the remainder are relatively small, and iii) the values of $E(x_1) = \theta_1$, $E(x_2) = \theta_2$, $\mathrm{Var}(x_1)$, $\mathrm{Var}(x_2)$ and $\mathrm{Cov}(x_1, x_2)$, are implicitly assumed to be known when in practice they are unknown and must be estimated from the data. How good an approximation equation 2-16 yields is therefore dependent on the quality (precision and accuracy) of the estimates used in the formula. This in turn depends on the size of the data set, whether the data are balanced or unbalanced, and the estimation procedure used. Factors affecting these estimates have been discussed previously.

Summary of the Assumptions Required

1. Variance of variance components

   i) ANOVA-based estimates

      Normality assumed: all random effects in the model are normally distributed with zero mean, finite and constant variance, and zero covariance between all effects in the model.

      If data are unbalanced, it must be further assumed that the sums of squares are distributed as multiples of central chi-squared distributions, when in fact they are not.

   ii) REML-based estimates

      Normality assumed, as for ANOVA-based estimates

      The sample size is sufficiently large such that the asymptotic variances and covariances are approximately correct.

2. Variance of a ratio (x_1/x_2)

   i) Dickerson's method

      Denominator (x_2) assumed to be known without error, i.e., it is a known constant.

   ii) Taylor series approximation

      The remainder term is assumed to be small, i.e., all second order and higher terms are assumed to be negligible and all covariances between the x_i's are assumed to be negligible.

      $E(x_1)$, $E(x_2)$, $Var(x_1)$, $Var(x_2)$, $Cov(x_1, x_2)$ are assumed known, but must be estimated from the data.

Case Study: Heritability Estimates

Two types of heritability estimates are commonly reported in the forestry literature: i) biased heritability (Ades *et al.* 1992, Borralho *et al.* 1992, Carson 1989, Cotterill *et al.* 1987, Dean *et al.* 1986, Dieters *et al.* 1990, Johnson and Burdon 1990, King *et al.* 1988, Sluder 1993, Woolaston *et al.* 1990), and ii) unbiased heritability (Balocchi *et al.* 1993, Barnes *et al.* 1992, Hodge and White 1992, Kiss and Yeh 1988, Lowe and Greene 1990, White and Hodge 1992, Yeh and Heaman 1982). Biased estimates of heritability are obtained when variance components are estimated from data collected on single sites, while unbiased heritability estimates result from analyses combining data from paired or multiple test locations. In the analysis of single-site data the interaction between additive genetic effects and environmental effects is confounded with additive genetic effects. The inability to separate additive-by-environment effects from additive effects means that the estimated additive genetic variance, and consequently the heritability, is biased upwards (Dickerson 1962, Comstock and Moll 1963).

Materials and Methods

Experimental material

Variance components were estimated from single-site analyses of 148 separate full-sib progeny tests of slash pine established between 1966 and 1988, on sites located primarily in Florida, Georgia, and Alabama, with a few tests in Mississippi and South Carolina of the USA. Each test was established in a randomized complete block design, with each full-sib family being represented by either a single row plot or non-contiguous plot composed of between 5 and 10 trees in each block. There were between 6 and 30 parents (mean of 16.4) in each test, crossed in either a diallel or factorial mating design, forming between 10 and 57 full-sib families (mean

of 29.2). For the single-site analyses (to estimate h_b^2), the tests were split into one of two size groups based on the number of parents in the test: 'small' (6-15 parents) and 'large' (16-30 parents).

Measurements of total height, diameter (4.5 ft or 1.37 m), and the presence/absence of fusiform rust infection were collected at between four and twelve years after planting. Measurements were grouped into three age classes centered at five, eight, and eleven years. Tree volume (Goddard and Strickland 1968, Rockwood 1981) and a rust score, 0 for no rust and 100 if infected with rust, were subsequently calculated for each tree.

As a consequence of the tests being established over a 23-year period, not all tests were measured at all three ages (Table 2-1). For this study, data were available for 121, 100, and 81 tests at five, eight, and eleven years, respectively. The overall level of rust infection varied from 0 to near 100% infection. It is known that tests with low or high levels of rust provide little information on rust resistance (White and Hodge 1989); therefore, only tests with more than 20% and less than 85% rust were included. Hence the study includes only 113 tests assessed for rust infection, with 78, 70, and 59 tests at five, eight, and eleven years respectively.

From these 148 full-sib slash pine tests, it was possible to identify 76 independent pairs of connected tests, with no tests represented in more than one test pair. However, only 39 of these test pairs had a rust infection level greater that 20% and less than 85% in both tests of the pair. These test pairs were used to obtain independent estimates of unbiased heritability. The much smaller number of test pairs than individual tests did not make it possible to divide the test pairs into groups based on test size as was done in case of biased heritability.

Table 2-1. Overall test means for volume and fusiform rust resistance in full-sib progeny tests of slash pine which were used in the case study of the variance of heritability estimates.

Trait	Age Class (Yrs)	Number of Tests	Volume (ft^3)			Rust (%)		
			Mean	Min	Max	Mean	Min	Max
Volume	5	121	0.2220	0.0443	0.7958	28.8	0.0	81.9
	8	100	1.1037	0.2814	2.9868	33.4	0.4	92.7
	11	81	2.0335	0.6938	4.6277	35.9	2.1	90.8
Rust	5	77	0.2391	0.0552	0.7958	40.2	20.2	81.9
	8	70	1.1066	0.2814	2.4610	41.4	20.3	83.5
	11	60	2.1176	0.6938	4.6277	43.9	20.1	81.7

Estimation of variance components

Variance components were estimated using DIALL (Schaffer and Usanis 1969) for single-site analyses, and GAREML (Huber 1993) for both single- and paired-site analyses. Prior to conducting the paired-site analyses, scale effects were removed by dividing the volume of each tree by the square root of the within plot variance (estimated using REML on the individual site data). The first program yields Henderson's Method III (HM3), ANOVA-based estimates, while GAREML provides REML estimates using Geisbrecht's (1983) algorithm. Negative estimates obtained using HM3 were accepted to ensure that these estimates were unbiased; however, REML estimates are by definition nonnegative.

The most complex model used was that for paired-site analyses of factorial tests, using GAREML, while all other models can be thought of as a subset of this model. This model was

$$y_{ijklmo} = \mu + t_i + b_{ij} + set_o + f_k + m_l + tf_{ik} + tm_{il} + fm_{kl} + tfm_{ikl} + p_{ijkl} + e_{ijklm}$$

where y_{ijklm} is the m^{th} tree in the kl^{th} family in the o^{th} set, and j^{th} block of the i^{th} test,

μ is the population mean,

t_i is the random effect of the i^{th} test environment, $E(t_i)=0$ and $Var(t_i)=\sigma_t^2$,

b_{ij} is the random effect of the j^{th} block in the i^{th} test, $E(b_{ij})=0$ and $Var(b_{ij})=\sigma^2_b$,

set_o is the random effect of the o^{th} disconnected set, $E(set_o)=0$ and $Var(set_o)=\sigma^2_s$,

f_k is the random effect of the k^{th} female, $E(f_k)=0$ and $Var(f_k)=\sigma^2_{gca}$,

m_l is the random effect of the l^{th} male, $E(m_l)=0$ and $Var(m_l)=\sigma^2_{gca}$,

fm_{kl} is the random effect of the interaction between the k^{th} female and the l^{th} male, $E(fm_{kl})=0$ and $Var(fm_{kl})=\sigma^2_{sca}$,

tf_{ik} is the random interaction between the i^{th} test and k^{th} female, $E(tf_{ik})=0$, and $Var(tf_{ik})=\sigma^2_{tgca}$,

tm_{il} is the random interaction between the i^{th} test and l^{th} male, $E(tm_{il})=0$, and $Var(tm_{il})=\sigma^2_{tgca}$,

tfm_{ikl} is the random interaction between the i^{th} test and kl^{th} family, $E(tfm_{ikl})=0$, and $Var(tfm_{ikl})=\sigma^2_{tsca}$,

p_{ijkl} is the random effect of the $ijkl^{th}$ plot, $E(p_{ijkl})=0$ and $Var(p_{ijkl})=\sigma^2_p$, and

e_{ijklm} is the random effect within the $ijkl^{th}$ plot, $E(e_{ijklm})=0$ and $Var(e_{ijklm})=\sigma^2_w$.

In this model it was assumed that there was no covariance between the random effects, that variances due to the female and male effects were equal, and that the female and male-by-environment interactions are the same. GAREML produced one estimate of σ^2_{gca} and σ^2_{tgca} by pooling the estimates from the male and female parents.

In the case of diallel tests, individual parents are used as both males and females in the crosses. By assuming the absence of any reciprocal effects, i.e., that it does not matter whether a parent is used as a male or as a female, it is possible to estimate the variance components for GCA and GCA-by-test location. Thus, in the above linear model f_k, m_l, tf_{ik}, and tm_{il} were replaced by g_k, g_l, tg_{ik}, and tg_{il}, respectively, for the general combining ability of the k^{th} (or l^{th}) parent, and fm_{kl} was replaced by s_{kl} for the specific combining ability (Griffing 1956).

When conducting single-site analyses, the models used for factorial and diallel experiments were as given above, except that all terms involving effects due to the i^{th} test were dropped from the model. Further, the program DIALL does not allow the inclusion of a term for disconnected sets or of a term for the variation between plots (Schaffer and Usanis 1969); therefore, in this case any plot effect was confounded with the within plot error. Thus, for HM3 estimates from DIALL the expectation of the within plot error is $E(\hat{\sigma}^2_W) = \sigma^2_p + \sigma^2_w$.

Estimation of biased heritability and its variance

Estimates of narrow sense, biased heritability were obtained using the formulae

i) HM3,

$$\hat{h}^2_b = \frac{4\hat{\sigma}^2_{GCA}}{2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_W}$$

ii) REML,

$$\hat{h}^2_b = \frac{4\hat{\sigma}^2_{GCA}}{2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_p + \hat{\sigma}^2_w}$$

where $\hat{\sigma}^2_{GCA}$ and $\hat{\sigma}^2_{SCA}$ when estimated from single sites are confounded with genotype-by-environment interaction, thus $E(\hat{\sigma}^2_{GCA}) = \sigma^2_{gca} + \sigma^2_{tgca}$ and $E(\hat{\sigma}^2_{SCA}) = \sigma^2_{sca} + \sigma^2_{tsca}$ (Comstock and Moll 1963). These formulae for biased heritability, are similar to those used in forest genetics (Cotterill 1987), and are appropriate for use with data corrected for set and block effects.

Estimates of the variance of biased heritability estimates were obtained for both the HM3 and REML variance component estimates, using i) Dickerson's method, ii) the Taylor series approximation, and iii) an empirical estimate of the variance. The empirical estimate of the variance of heritability is simply the variance among comparable heritability estimates. This

empirical estimate therefore contains an estimate of the variance of heritability estimates due to the differences in site quality and uniformity, differences in management and maintenance of the tests, differences in the mating design and experimental design, as well as many other factors. It is believed, given the large number of different progeny tests involved in this empirical estimate, that the empirical estimate approached the 'true' variance of heritability estimates, of slash pine in the southeastern USA at these ages.

The formulae used for calculating the Dickerson's approximate estimate of the variance of the biased heritability were

i) HM3,

$$\hat{Var}(h^2_b) \cong \frac{16\,\hat{Var}(\hat{\sigma}^2_{GCA})}{(2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_{W})^2}$$

ii) REML,

$$\hat{Var}(h^2_b) \cong \frac{16\,\hat{Var}(\hat{\sigma}^2_{GCA})}{(2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{W})^2}$$

The formulae used for calculating the variance of the biased heritability, estimated using the Taylor series approximation, were obtained by inserting into equation 2-16 the approximate values of $E(x_1)=\theta_1$, $E(x_2)=\theta_2$, $Var(x_1)$, $Var(x_2)$, and $Cov(x_1, x_2)$ given below:

i) HM3,

$$E(x_1) \cong 4\hat{\sigma}^2_{GCA}$$

$$E(x_2) \cong 2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_{W}$$

$$Var(x_1) \cong 16\hat{Var}(\hat{\sigma}^2_{GCA})$$

$$\begin{aligned}\hat{Var}(x_2) \cong\ & 4\hat{Var}(\hat{\sigma}^2_{GCA}) + \hat{Var}(\hat{\sigma}^2_{SCA}) + \hat{Var}(\hat{\sigma}^2_{W}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{SCA}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{W} + \\ & 2\hat{Cov}(\hat{\sigma}^2_{SCA}, \hat{\sigma}^2_{W})\end{aligned}$$

$$Cov(x_1, x_2) \cong 8\hat{Var}(\hat{\sigma}^2_{GCA}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{SCA}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{W})$$

ii) REML,

$$E(x_1) \cong 4\hat{\sigma}^2_{GCA}$$

$$E(x_2) \cong 2\hat{\sigma}^2_{GCA} + \hat{\sigma}^2_{SCA} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{w}$$

$$Var(x_1) \cong 16\hat{Var}(\hat{\sigma}^2_{GCA})$$

$$\begin{aligned}Var(x_2) \cong\ & 4\hat{Var}(\hat{\sigma}^2_{GCA}) + \hat{Var}(\hat{\sigma}^2_{SCA}) + \hat{Var}(\hat{\sigma}^2_{p}) + \hat{Var}(\hat{\sigma}^2_{w}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{SCA}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{p}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{w}) \\ & 2\hat{Cov}(\hat{\sigma}^2_{SCA}, \hat{\sigma}^2_{p}) + 2\hat{Cov}(\hat{\sigma}^2_{SCA}, \hat{\sigma}^2_{w}) + 2Cov(\hat{\sigma}^2_{p}, \hat{\sigma}^2_{w})\end{aligned}$$

$$\begin{aligned}Cov(x_1, x_2) \cong\ & 8\hat{Var}(\hat{\sigma}^2_{GCA}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{SCA}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{p}) + 4\hat{Cov}(\hat{\sigma}^2_{GCA}, \hat{\sigma}^2_{w})\end{aligned}$$

Estimation of unbiased heritability and its variance

Variance component estimates from the paired-site analyses were substituted into the following formula to calculate heritability estimates:

$$\hat{h}^2 = \frac{4\hat{\sigma}^2_{gca}}{2\hat{\sigma}^2_{gca} + \hat{\sigma}^2_{sca} + 2\hat{\sigma}^2_{tgca} + \hat{\sigma}^2_{tsca} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{w}}$$

The formula used for calculating the Dickerson's approximate estimate of the variance of the biased heritability was

$$Var(\hat{h}^2) \cong \frac{16\,\hat{Var}(\hat{\sigma}^2_{gca})}{(2\hat{\sigma}^2_{gca} + \hat{\sigma}^2_{sca} + 2\hat{\sigma}^2_{tgca} + \hat{\sigma}^2_{tsca} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{w})^2}$$

The formula used for calculating the Taylor series approximate variance of the heritability was obtained by inserting into equation 2-16 the approximate values of $E(x_1)=\theta_1$, $E(x_2)=\theta_2$, $Var(x_1)$, $Var(x_2)$, and $Cov(x_1, x_2)$ given below:

$$E(x_1) \cong 4\hat{\sigma}^2_{gca}$$

$$E(x_2) \cong 2\hat{\sigma}^2_{gca} + \hat{\sigma}^2_{sca} + 2\hat{\sigma}^2_{tgca} + \hat{\sigma}^2_{tsca} + \hat{\sigma}^2_{p} + \hat{\sigma}^2_{w}$$

$$Var(x_1) \cong 16\,\hat{Var}(\hat{\sigma}^2_{gca})$$

$$\begin{aligned} Var(x_2) \cong\ & 4\hat{Var}(\hat{\sigma}^2_{gca}) + \hat{Var}(\hat{\sigma}^2_{sca}) + 4\hat{Var}(\hat{\sigma}^2_{tgca}) + \hat{Var}(\hat{\sigma}^2_{tsca}) + \\ & \hat{Var}(\hat{\sigma}^2_{p}) + \hat{Var}(\hat{\sigma}^2_{w}) + 4\hat{Cov}(\hat{\sigma}^2_{gca}, \hat{\sigma}^2_{sca}) + \\ & 8\hat{Cov}(\hat{\sigma}^2_{gca}, \hat{\sigma}^2_{tgca}) + 4\hat{Cov}(\hat{\sigma}^2_{gca}, \hat{\sigma}^2_{tsca}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{gca}, \hat{\sigma}^2_{p}) + 4\hat{Cov}(\hat{\sigma}^2_{gca}, \hat{\sigma}^2_{w}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{sca}, \hat{\sigma}^2_{tgca}) + 2\hat{Cov}(\hat{\sigma}^2_{sca}, \hat{\sigma}^2_{tsca}) + \\ & 2\hat{Cov}(\hat{\sigma}^2_{sca}, \hat{\sigma}^2_{p}) + 2\hat{Cov}(\hat{\sigma}^2_{sca}, \hat{\sigma}^2_{w}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{tgca}, \hat{\sigma}^2_{tsca}) + 4\hat{Cov}(\hat{\sigma}^2_{tgca}, \hat{\sigma}^2_{p}) + \\ & 4\hat{Cov}(\hat{\sigma}^2_{tgca}, \hat{\sigma}^2_{w}) + 2\hat{Cov}(\hat{\sigma}^2_{tsca}, \hat{\sigma}^2_{p}) + \\ & 2\hat{Cov}(\hat{\sigma}^2_{tsca}, \hat{\sigma}^2_{w}) + 2\hat{Cov}(\hat{\sigma}^2_{p}, \hat{\sigma}^2_{w}) \end{aligned}$$

Results and Discussion

Distribution of heritability estimates

The frequency distributions of the biased heritability estimates obtained using variance components from HM3 and REML estimation are plotted in Figure 2-2. Figures 2-2a), 2-2b), and

2-2c) are the frequency distributions for the estimated heritability of volume at 5, 8, and 11 years respectively. The distributions of the rust heritability estimates are plotted in Figures 2-2d), 2-2e) and 2-2f), for 5, 8, and 11 years, respectively.

The biased heritability estimates obtained using HM3 variance component estimates all included negative estimates of the heritability, while those which use REML estimates were truncated at zero. This reflected the differences in the two methods of variance component estimation — HM3 allowed negative estimates, while the REML algorithm used sets the negative estimates to zero and revaluated the system of equations. In general, HM3 and REML yielded similar frequency distributions, except that estimates which were negative in HM3, were concentrated at, or just above, zero when using REML (Figure 2-2). The distributions for the REML estimates also appear to be smoother than those from HM3.

The only exception to these observations is volume at five years. In this case there was a marked disparity between the frequency distribution of the HM3 and REML estimates of biased heritability. The reason for this marked difference between the HM3 and REML estimates is difficult to determine. However, it seems possible that it may reflect the small absolute size of the volumes measured at 5 years (Table 2-1). In the case of HM3 estimates, the small absolute size of the observations led to small sums of squares and consequently a high probability of variance component estimates being close to zero, or negative. Conversely, in the case of REML, the estimation procedure tended to give zero variance components for factors with very small amounts of associated variance, especially if they were at all unstable within the parameter space. For five-year volume, this often resulted in estimates for all variance components, except σ^2_{gca} and σ^2_w, being small or zero, which then tended to produce larger estimates of h^2_b than those estimated with HM3.

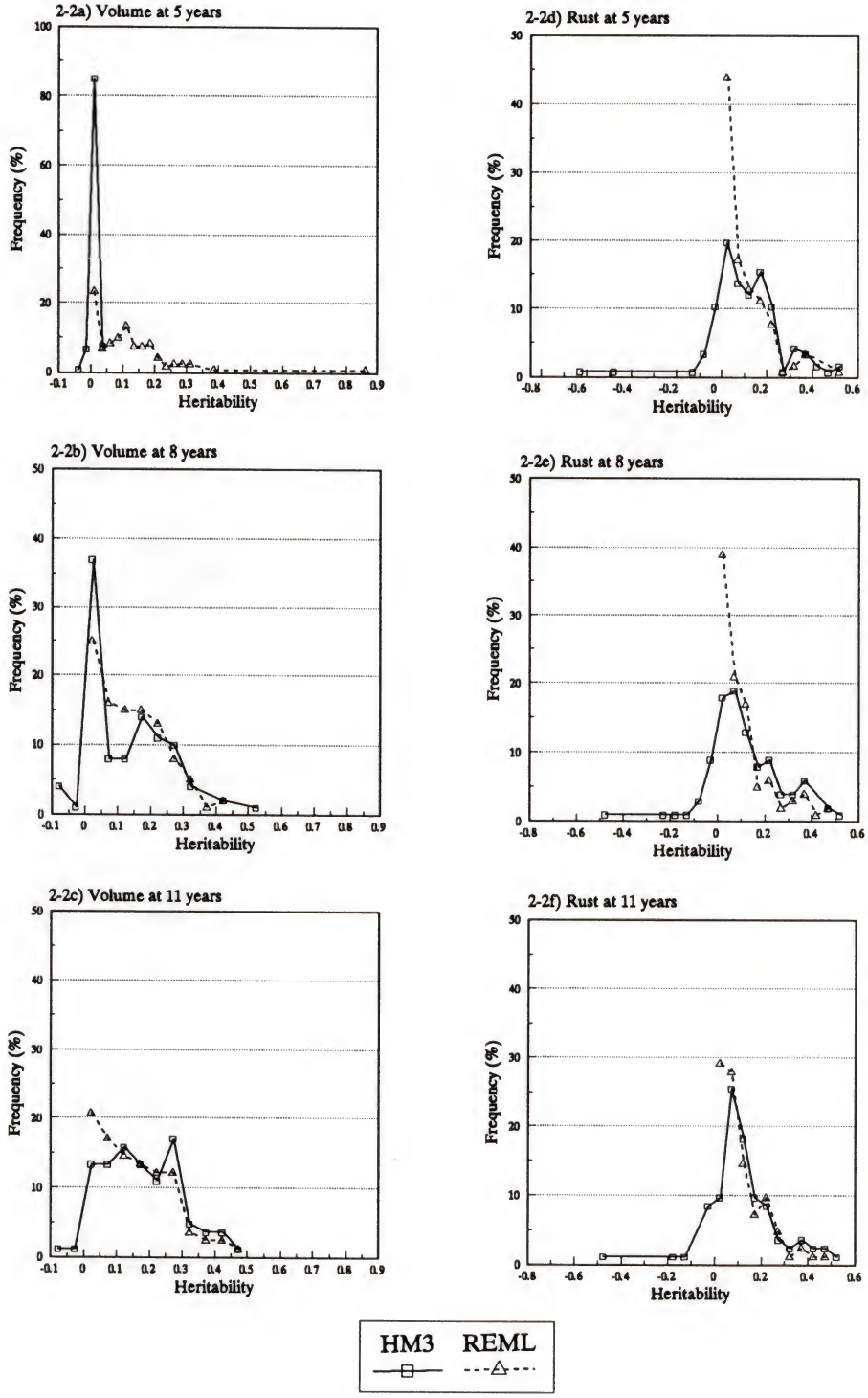

Figure 2-2. Frequency distribution of biased heritability estimates from single-site analyses of slash pine, using Henderson's Method III (HM3) and Restricted Maximum Likelihood (REML) to estimate the variance components. Figures 2-2a), 2-2b) and 2-2c) are for volume at 5, 8, and 11 years respectively, while Figures 2-2d), 2-2e) and 2-2f) are for rust at 5, 8, and 11 years respectively.

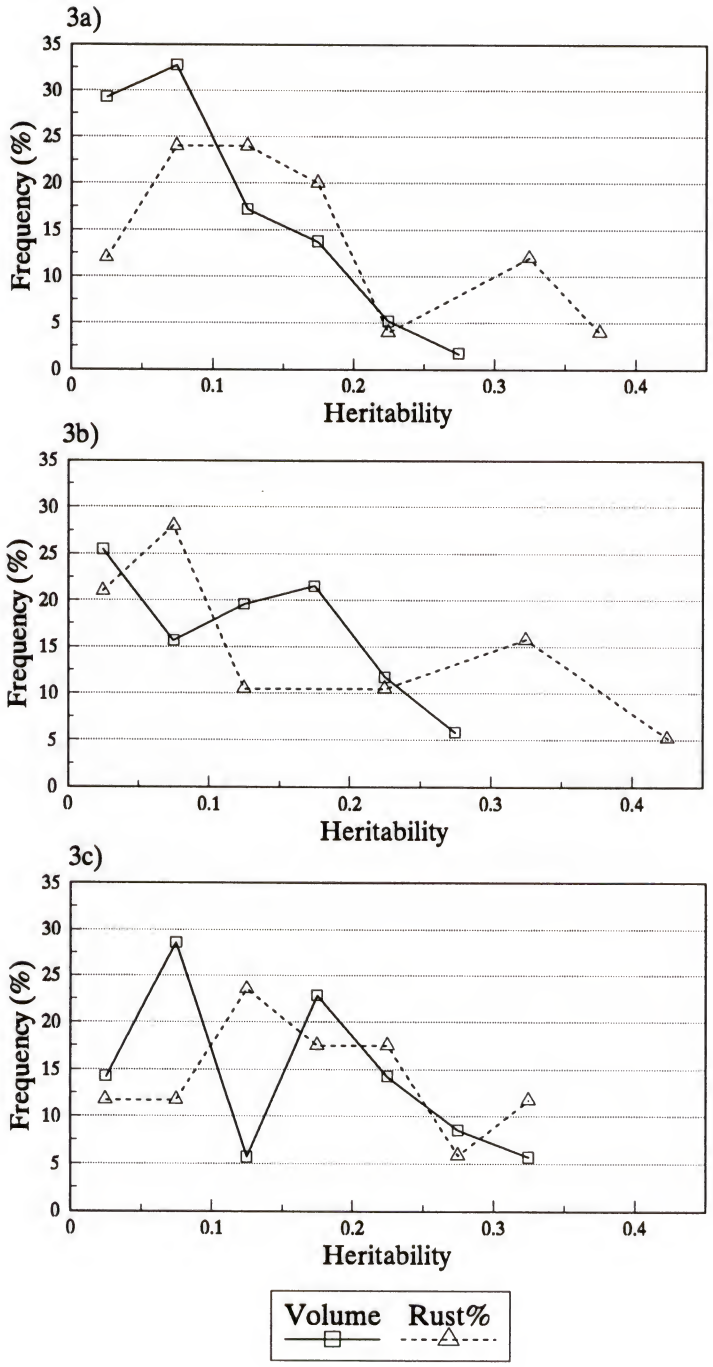

Figure 2-3. Frequency distribution of unbiased heritability estimates for volume growth and rust percent, estimated in paired full-sib tests of slash pine at a) 5 years, b) 8 years, and c) 11 years. All variance component estimates obtained using Restricted Maximum Likelihood (REML).

The distributions of unbiased heritability estimates (Figure 2-3) were similar to those of the biased heritability estimates (Figure 2-2). However there was considerably more fluctuation in the distribution of estimates; probably due to the lower number of observations which contributed to these graphs. The distributions of unbiased heritability generally do not peak in the smallest size class, which always occurred in the biased heritabilities estimated using REML (Figure 2-2), indicating a lower incidence of zero estimates for σ^2_{gca} from paired tests than from individual tests.

The existence of considerable variation between individual heritability estimates is clearly demonstrated by Figures 2-2 and 2-3, regardless of whether HM3 or REML was used to estimate the variance components. Although the data presented are only from slash pine, there is no reason to assume that this situation is unique to this species. Therefore it becomes apparent that individual heritability estimates are unreliable.

Dickerson's versus Taylor series approximations

The variances of the biased heritability estimates of volume and rust resistance, as estimated by the Dickerson and Taylor series approximations, were compared using simple linear regression (Tables 2-2 and 2-3). This analysis revealed a very strong linear relationship between the two approximation methods in both traits. All R-squares were greater than 0.92, and the strength of this relationship was apparently unaffected by the method used to estimate the variance components. For volume, the intercept terms were never significantly different from zero for either HM3 or REML estimates. However for rust resistance when using HM3 estimates the intercept was significantly different from zero in all cases except for large tests at 11 years. For REML estimates of rust resistance the intercept was significant only for small tests at 11 years. All slopes were estimated to be greater than one, except in the case of rust in small tests when using HM3. Slope parameters are significant at $p=0.0001$ in all models.

For tree volume using either HM3 or REML estimates of variance components, and rust resistance using REML estimate, it appears that Dickerson's approximation was in general slightly more conservative than the Taylor series approximation of the variance of the heritability. For HM3 estimates of rust resistance, in general the intercept was significantly greater than zero even though slopes were less than one in the small tests, which also suggested that Dickerson's method was more conservative.

The overall means for biased heritability in each age group (volume in Tables 2-4 and 2-5 and rust in Tables 2-6 and 2-7) showed that the mean variance estimated by Dickerson's method was usually slightly larger, but never smaller, than that estimated by the Taylor series approximation. In the case of the estimated heritability of rust obtained in the small tests using HM3 variance components (Table 2-6), the means of Dickerson's approximation were relatively small when compared to their standard errors, which may account for the slopes being less than 1.0 (Table 2-3), even though the overall means always showed Dickerson's method to be slightly conservative when compared to the Taylor series approximation.

A comparison of the Dickerson and Taylor series approximate variances of the unbiased heritability estimates for volume and rust resistance verified the results with the biased heritability estimates (Tables 2-8 and 2-9). In all cases the slopes of the simple linear regression analyses between the Dickerson and Taylor series estimates were greater than 1.0 and the R-squares were always greater than 0.98 (Table 2-8); however, except for 8 year volume and 5 year rust, the intercept term was significantly less than zero. Nevertheless, when we examine the overall means of Dickerson and Taylor series approximations (Table 2-9), the Dickerson method yielded estimates which were on average slightly larger than those from the Taylor series approximation. Thus for both biased and unbiased estimates of heritability, there was a strong linear relationship

between the estimates of the variance of heritability estimates obtained by the Dickerson and Taylor series, and Dickerson's method generally appeared to be slightly more conservative.

Table 2-2. Regression coefficients and R-squares from simple linear regression of Dickerson's versus Taylor series estimates of the variance of biased heritability (h_b^2) estimates for volume in slash pine for small and large experiments. (Model: Dickerson's = Intercept + Taylor * slope; all models are significant at the 0.0001 level)

Age Class	Size Class	# Tests	Henderson's Method 3			REML		
			Int.	Slope	R^2	Int.	Slope	R^2
Five years	Small	60	0.0000	1.0175	0.9997	-0.0004	1.1955	0.9828
	Large	61	0.0000	1.0089	0.9995	-0.0002	1.1780	0.9805
Eight years	Small	46	-0.0008	1.3276	0.9448	-0.0007	1.2557	0.9680
	Large	52	0.0001	1.1587	0.9610	-0.0006	1.2576	0.9656
Eleven years	Small	43	-0.0003	1.2701	0.9810	-0.0005	1.2722	0.9817
	Large	38	0.0001	1.1774	0.9197	-0.0002	1.2126	0.9510

Table 2-3. Regression coefficients and R-squares from simple linear regression of Dickerson's versus Taylor series estimates of the variance of biased heritability (h_b^2) estimates for rust resistance in slash pine for small and large experiments. (Model: Dickerson's = Intercept + Taylor * slope; all models are significant at the 0.0001 level)

Age Class	Size Class	# Tests	Henderson's Method 3			REML		
			Int.	Slope	R^2	Int.	Slope	R^2
Five years	Small	32	0.0031	0.9456	0.9976	0.0008	1.0731	0.9731
	Large	46	0.0017	1.0389	0.9583	-0.0006	1.0970	0.9534
Eight years	Small	33	0.0031	0.9634	0.9855	0.0005	1.1216	0.9739
	Large	37	-0.0016	1.5325	0.9469	-0.0017	1.5581	0.9748
Eleven years	Small	27	0.0017	0.9480	0.9980	0.0008	1.2787	0.9699
	Large	32	0.0011	1.0979	0.9296	0.0003	1.1610	0.9492

Table 2-4. Overall mean biased heritability (h_b^2) of single-site volume in slash pine estimated using Henderson's Method 3 (HM3), and the estimated mean variance (± standard error) of h_b^2 as estimated using the Dickerson's and Taylor series approximations, and an empirical estimate, for three age classes and three size classes.

Age Class	Size Class	# Tests	Mean h_b^2	Estimated Mean Variance of Heritability		
				Dickerson	Taylor	Empirical
Five years	Small	60	0.0048	0.00001±0.000006	0.00001±0.000006	0.00006
	Large	61	0.0073	0.00004±0.000010	0.00004±0.000010	0.00001
	All	121	0.0061	0.00003±0.000006	0.00003±0.000006	0.00010
Eight years	Small	48	0.1169	0.00968±0.001712	0.00787±0.001253	0.01567
	Large	52	0.1274	0.00743±0.000947	0.00636±0.000801	0.01346
	All	100	0.1224	0.00851±0.000960	0.00709±0.000732	0.01440
Eleven years	Small	43	0.1770	0.01483±0.002114	0.01190±0.001648	0.01208
	Large	38	0.1809	0.01136±0.001110	0.00958±0.000904	0.01455
	All	81	0.1788	0.01320±0.001245	0.01081±0.000975	0.01308

Table 2-5. Overall mean biased heritability (h_b^2) of single-site volume in slash pine estimated using Restricted Maximum Likelihood (REML), and the estimated mean variance (± standard error) of h_b^2 as estimated using the Dickerson's and Taylor series approximations, and an empirical estimate, for three age classes and three size classes.

Age Class	Size Class	# Tests	Mean h_b^2	Estimated Mean Variance of Heritability		
				Dickerson	Taylor	Empirical
Five years	Small	60	0.1097	0.01093±0.001422	0.00944±0.001179	0.00688
	Large	61	0.1183	0.00656±0.000801	0.00571±0.000673	0.01820
	All	121	0.1140	0.00872±0.000833	0.00756±0.000694	0.01250
Eight years	Small	48	0.1397	0.01507±0.002063	0.01256±0.001621	0.00994
	Large	52	0.1407	0.01018±0.001248	0.00855±0.000975	0.01279
	All	100	0.1402	0.01253±0.001203	0.01047±0.000945	0.01131
Eleven years	Small	43	0.1709	0.01999±0.003048	0.01613±0.002374	0.01134
	Large	38	0.1433	0.01134±0.001607	0.00953±0.001293	0.01431
	All	81	0.1579	0.01593±0.001839	0.01303±0.001438	0.01276

Table 2-6. Overall mean biased heritability (h_b^2) of single-site fusiform rust resistance in slash pine estimated using Henderson's Method 3 (HM3), and the estimated mean variance (± standard error) of h_b^2 as estimated using the Dickerson's and Taylor series approximations, and an empirical estimate, for three age classes and three size classes.

Age Class	Size Class	# Tests	Mean h_b^2	Estimated Mean Variance of Heritability		
				Dickerson	Taylor	Empirical
Five years	Small	32	0.1131	0.03807±0.01653	0.03699±0.01746	0.04139
	Large	46	0.1858	0.01167±0.00253	0.00958±0.00239	0.01982
	All	78	0.1560	0.02250±0.00704	0.02083±0.00739	0.02954
Eight years	Small	33	0.1139	0.02648±0.00797	0.02425±0.00821	0.03120
	Large	37	0.1779	0.00902±0.00130	0.00692±0.00823	0.02256
	All	70	0.1477	0.01725±0.00393	0.01509±0.00400	0.02728
Eleven years	Small	27	0.1051	0.02043±0.00955	0.01977±0.01007	0.02544
	Large	32	0.1803	0.01207±0.00199	0.01000±0.00175	0.02025
	All	59	0.1459	0.01590±0.00449	0.01447±0.00470	0.02365

Table 2-7. Overall mean biased heritability (h_b^2) of single-site fusiform rust resistance in slash pine estimated using Restricted Maximum Likelihood (REML), and the estimated mean variance (± standard error) of h_b^2 as estimated using the Dickerson's and Taylor series approximations, and an empirical estimate, for three age classes and three size classes.

Age Class	Size Class	# Tests	Mean h_b^2	Estimated Mean Variance of Heritability		
				Dickerson	Taylor	Empirical
Five years	Small	32	0.1045	0.01289±0.00309	0.01126±0.00284	0.00993
	Large	46	0.1546	0.00988±0.00176	0.00830±0.00157	0.01195
	All	78	0.1340	0.01111±0.00163	0.00951±0.00149	0.01159
Eight years	Small	33	0.1004	0.01278±0.00366	0.01093±0.00322	0.00937
	Large	37	0.1677	0.00999±0.00180	0.00754±0.00114	0.01832
	All	70	0.1360	0.01131±0.00196	0.00913±0.00163	0.01505
Eleven years	Small	27	0.1268	0.00778±0.00172	0.00778±0.00132	0.00843
	Large	32	0.1405	0.00971±0.00191	0.00810±0.00161	0.01264
	All	59	0.1342	0.00953±0.00129	0.00795±0.00105	0.01058

Table 2-8. Regression coefficients and R-squares from simple linear regression of Dickerson's versus Taylor series estimates of the variance of heritability (h^2) estimates for volume and rust in slash pine. (Model: Dickerson's = Intercept + Taylor * slope; all models are significant at the 0.0001 level)

Trait	Age Class (Yrs)	Number of Tests	REML		
			Int.	Slope	R^2
Volume	5	62	-0.0004	1.2180	0.9882
	8	53	-0.0002	1.1655	0.9902
	11	36	-0.0010	1.3260	0.9905
Rust	5	28	0.0006	1.1103	0.9817
	8	21	-0.0008	1.3975	0.9860
	11	18	-0.0011	1.4206	0.9876

Table 2-9. Overall means for heritability of paired-site volume and rust in slash pine, estimated using Restricted Maximum Likelihood (REML). The mean heritability (h^2) and its mean variance ($\pm$ standard error) as estimated using the Dickerson's and Taylor series approximations, compared to an empirical estimate.

Trait	Age Class (Yrs)	# Tests	Mean h^2	Estimated Mean Variance of Heritability		
				Dickerson	Taylor	Empirical
Volume	5	62	0.0912	0.005181±0.000615	0.004617±0.000502	0.004849
	8	53	0.1181	0.007188±0.000741	0.006323±0.000633	0.006342
	11	36	0.1529	0.011918±0.000233	0.009759±0.001749	0.007657
Rust	5	28	0.1519	0.008621±0.002165	0.007243±0.001932	0.009479
	8	21	0.1471	0.007315±0.001875	0.005772±0.001332	0.015803
	11	18	0.1622	0.007438±0.001475	0.006041±0.001032	0.008535

Empirical estimate versus Dickerson and Taylor series approximations

The mean variance of h_b^2 for both volume growth and rust, estimated using the Dickerson, Taylor series and empirical methods (Tables 2-4, 2-5, 2-6 and 2-7) showed that the approximate

methods were always of a magnitude similar to the empirical variance estimate, regardless of the procedure used to estimate the variance components and the size of the test. However the empirical estimate of the variance of h_b^2 was usually larger than that estimated by either of the approximate methods. The more conservative nature of the Dickerson estimates means that these were generally closer to the empirical estimates than are the Taylor series estimates.

The only reasonably consistent exception to this observed underestimation of the empirical variance was when using REML estimates from the small tests. In this case the approximate variances were nearly the same as, or larger, than the empirical estimate of variance. It is possible that the difference between HM3 and REML in small tests reflected the fact that asymptotic variances and covariances were used for REML estimates. Therefore, the Taylor series and Dickerson approximations, which use the asymptotic variances and covariances, would have been affected by sample size. Thus, it might reasonably be expected that estimates from larger tests would be more reliable than those from small tests when using REML.

Unbiased heritability estimates also showed a good relationship between the empirical variance estimates and the Dickerson and Taylor series approximations. However for rust resistance the approximate methods were consistently low, but only significantly so at 8 years (Table 2-9). It is possible that the small number of test pairs (less than 30) which contributed to the rust resistance estimates in Table 2-9 may have been a contributing factor. Nevertheless, the estimates obtained from the two approximate methods were always of a similar magnitude to the empirical estimates, and there is no strong evidence to suggest that the approximate methods consistently over or underestimated the 'true' variance of paired test unbiased heritability estimates.

It is unlikely that all the assumptions required in the development and application the Dickerson and Taylor series approximations were fulfilled for both HM3 and REML variance

component estimates. Nevertheless these two methods provided reasonable approximations to the empirical variance of biased and unbiased heritabilities for both volume and rust resistance in slash pine. The observed deviations of both the Dickerson and Taylor series approximations from the empirical estimates of the variance only serves to highlight the fact that these are approximate and not exact methods.

Conclusions

The data presented indicate that Dickerson's method provided estimates of the variance of h_b^2 and h^2 that corresponded closely to estimates obtained from the Taylor series approximation, but which were slightly more conservative (i.e., larger). This conclusion was found to be valid for both tree volume (a continuous trait) and fusiform rust resistance (a bernoulli trait) in slash pine, and was not dependent on the method of variance component estimation or on test size.

Comparisons of the Dickerson's and Taylor series approximations with an empirical estimates of the variance of h_b^2 and h^2 indicated that these are both reasonable approximations to the 'true' variance (taken to be the empirical estimate) of h_b^2 and h^2. However these approximate variances tended to underestimate the 'true' variance of h_b^2, but for h^2 there is no strong evidence to suggest that the approximate methods consistently over or underestimated the empirical variance. When REML estimates from small tests (less than 16 parents in this case) are used, it is possible that the asymptotic variance and covariances may not be reliable, causing an over-estimation of the empirical variance of h_b^2.

For most practical purposes when seeking to estimate the variance of biased or unbiased heritability estimates, it seems reasonable to conclude that either the Dickerson or the Taylor series approximations will provide useful estimates of the 'true' variance. However Dickerson's

method seems to be preferable to the Taylor series approximation because i) both methods provide essentially the same information about the variance of heritability estimates, ii) it is more conservative than the Taylor series approximation and thus perhaps closer to the presumed 'true' variance, and iii) the Dickerson method is considerably simpler to calculate.

CHAPTER 3
GENETIC PARAMETER ESTIMATES FOR VOLUME FROM FULL-SIB TESTS OF SLASH PINE (*PINUS ELLIOTTII*)

Introduction

In the genetic improvement of forest trees, much effort is invested to determine accurate and precise estimates of genetic parameters such as heritability, genetic correlations, and genotype-by-environment interactions. These parameters are used throughout the breeding cycle (White 1987) and are key elements in the determination of breeding, propagation, and deployment strategies. The importance of such parameter estimates in tree improvement is attributable to both the broad utility of such estimates and the long generation interval of most forest tree species (usually 10 years or more). Generation intervals of this length require high quality information on the genetic parameters to help ensure that genetic gains are maximized per unit time.

Precise and accurate genetic parameter estimates for growth traits in slash pine (*Pinus elliottii* Engelm. var *elliottii*) have been obtained from open-pollinated (half-sib) progeny tests (Hodge and White 1992). Half-sib tests allow the estimation of general combining ability (GCA); however, specific combining ability (SCA, attributable to dominance) can not be estimated from such tests. There are few published estimates of dominance in slash pine, and these are based on a small number of tests (e.g. Cotterill *et al.* 1987 — one test; Kraus 1973 — one test; Pswarayi 1993 — three tests).

The estimation of dominance variance requires a systematic mating design, with full-sib and half-sib progeny groups. Such tests are difficult, expensive and time consuming to establish,

and frequently were established as much as ten years after half-sib progeny tests. The lack of adequate information on dominance variance has commonly led tree breeders to assume that the amount of dominance is relatively small compared to additive variance. However, if this assumption is not true, it is possible that breeding programs have foregone significant additional genetic gain because their strategies did not allow for the possibility of capturing and utilizing genetic variation arising from dominance. In large-scale intensive plantation programs this additional gain, even if small, may have substantial economic value.

The Cooperative Forest Genetic Research Program (CFGRP), a cooperative tree improvement program operating in the southeastern United States, has maintained an intensive breeding program with slash pine since the early 1950's, and provides genetically improved material for a plantation estate of approximately 4 million hectares. An integral part of this program has been the establishment of many (a total of 265 tests [Hodge *et al.* 1991]) full-sib progeny tests of first-generation parents. These full-sib tests provide an extensive data set which can i) serve as an independent data set to validate genetic parameter estimates previously developed from open-pollinated tests (Hodge and White 1992), and ii) provide the first reliable estimates of the importance of dominance variance in slash pine.

Materials and Methods

Full-Sib Progeny Test Data

During the period 1956-1963 mass selection in natural stands identified over 2500 first-generation slash pine trees, most of which were subsequently grafted into multiclonal seed orchards (White *et al.* 1986). By the early 1970's it had become evident that relatively few of the original first-generation parents were resistant to infection by fusiform rust, *Cronartium fusiforme* Berk. Miyabe ex. Shirai f. sp. *fusiforme* (Goddard 1980), and 550 new first-generation trees free

of fusiform rust (White *et al.* 1987) were selected in stands between 10 and 20 years of age, where the average rust infection level exceeded 70% (Hendrickson 1976).

The full-sib tests established by the CFGRP comprise two separate test series. The first series, established between the period 1966-1973, utilized a factorial crossing scheme amongst some of the original 2500 first-generation selections (a total of 78 tests). The second series of tests was established between 1975 and 1989 (187 tests in total), utilizing a) the better first-generation parents in diallels or factorials, or b) the rust-free selections (as male parents) and the better first-generation parents (as females) in factorial mating designs.

In April 1994, 171 different full-sib tests more than three years old were included in this study (Table 3-1). These tests are mainly in Florida, Georgia, and Alabama, but a few tests are in Mississippi and South Carolina. All 171 tests were established in a randomized complete block design with 3-12 blocks (mean=5.4 blocks). In each block, a full-sib family was represented by one plot of 5-10 trees (mean=6.9 trees) arranged as either a row-plot or a noncontiguous plot. In any one test there were between 6 and 86 full-sib families (mean=30.6 families), which were derived from crosses among 6 to 47 parents (mean=18.2 parents). In total the tests included progeny from over 700 parents (165 rust-free parents), represented by over 2100 different full-sib families (550 families involving rust-free parents), and approximately 170,000 individual trees.

At each measurement age (nominally 5, 8 and 12 growing seasons after planting), all trees were assessed for diameter at breast height (4.5 ft), height, and the presence/absence of rust infection. From these measurements individual tree volume was calculated (Goddard and Strickland 1968, Rockwood 1981), and each tree was given a rust score (0=no rust, and 100=infected with rust). To facilitate analyses, data were grouped into four age classes, centered at 5, 8, 11 and 14 years (Table 3-1).

The site index of each test was estimated using the mean height of the largest 43.5% of the trees (ranked on diameter) to estimate dominant height (Bailey and Brooks 1994). This dominant height was then used to estimate site index according to formulae developed by Pienaar *et al.* (1990). Analyses revealed that data less than eight years of age over-estimated site-index, while data older than eight years provided site index estimates not significantly different from those obtained with eight year data; thus, the estimated site index was adjusted to that estimated from eight year data. Using this 8-year adjusted site index (predicted at a base age of 25 years), tests were assigned to site index classes: low, medium and high. The low and the high classes were composed of one quartile (lower or upper), and the medium class the middle two quartiles. The range of site indexes in each group were 43-63 ft (13.1-19.3 m), 64-75 ft (19.5-23 m), and 76-101 ft (23.2-31 m), for the low, medium and high groups respectively.

Variance component estimation

Considerable imbalance exists within this set of full-sib progeny tests, as commonly occurs in full-sib tests of long-lived perennial species. Specifically, i) the number of parents contributing progeny to any one test ranges from 6 to 47; ii) the number of parents in common between any given pair of tests ranges between 1 and 45; and, iii) in any given test or pair of tests, the number of crosses per parent usually were not equal. There is also further imbalance at the plot and individual levels which result from mortality during the life of the tests (Table 3-1). Consequently, traditional ANOVA-based estimators of variance components, such as Henderson's method III, are inefficient for use here (Searle *et al.* 1992).

GAREML (Huber 1993), a computer program which utilizes Giesbrecht's (1983) algorithm to provide REML (restricted maximum likelihood) estimates (Patterson and Thompson 1971) of variance components, has been demonstrated to provide variance component estimates with desirable properties when applied to data sets with the level and type of imbalance

commonly encountered in forest genetics tests (Huber 1993, p.82). REML estimation is also know to be relatively robust to violations of the underlying normality assumptions (Banks *et al.* 1985; Westfall 1987). On the basis of this work, REML estimation from individual tree data, using the program GAREML, was selected for use to estimate variance components in these full-sib tests of slash pine.

Table 3-1. Summary statistics of full-sib slash pine progeny tests, tabulated by age class. (Minimum, maximum and mean of overall test volume, rust infection and survival. Site index standardized to estimates from 8 year data.)

Age Class	Number of Tests	Summary Parameter	Age (yrs)	Volume per Tree (ft^3)	Rust Infection (%)	Survival (%)	Site Index (ft)
5 Years	142	Minimum	4	0.0443	0.00	41.87	43.9
		Maximum	6	0.7985	99.78	98.52	100.9
		Mean (se)	5.01 (0.01)	0.2212 (0.0104)	30.43 (1.87)	82.93 (1.13)	70.9 (0.78)
8 Years	118	Minimum	7	0.2814	0.36	35.72	43.9
		Maximum	9	2.9868	92.77	97.78	93.01
		Mean (se)	8.00 (0.08)	1.1274 (0.0466)	33.75 (2.00)	78.63 (1.32)	70.2 (0.83)
11 Years	85	Minimum	10	0.6938	2.07	35.45	43.9
		Maximum	12	5.1498	95.80	97.30	83.3
		Mean (se)	10.61 (0.09)	2.1519 (0.1172)	39.03 (2.61)	72.26 (1.61)	67.1 (0.87)
14 Years	21	Minimum	13	1.7134	14.79	53.36	56.8
		Maximum	15	5.9725	61.17	90.49	88.1
		Mean (se)	13.81 (0.19)	3.3231 (0.2394)	38.50 (2.69)	73.99 (2.31)	69.2 (1.76)

REML estimation using individual tree observations is numerically intense, and computationally expensive, making it necessary to break the problem of variance component estimation down into smaller, more manageable pieces. To achieve this, each full-sib test was

first analyzed separately, and to remove scale effects individual tree volumes were standardized by dividing the individual tree volumes by the square root of the within plot variance (σ_w) estimated from that test. Hence the within plot variance of the standardized data is equal to 1.00. Using the standardized data, all pairs of tests with at least five common parents were analyzed using GAREML in a pooled analysis. All possible connected test pairs and all combinations of measurement ages were analyzed in this manner (over 3000 separate paired analyses). Pairs were always constructed from data collected in different tests, and the same trees measured at different ages were never used in a paired analysis. Thus the potential problem of lack of independence of error terms (Hodge and White 1992) due to environmental correlation within a single site was avoided in these analyses.

When two tests have the same age, the REML estimates are variance components, but when the pair of tests are different ages, the estimates are considered covariance components between volume at different ages. This approach forces covariance component estimates to have the same properties as variance components, i.e., to be nonnegative. Previous estimates of additive genetic correlations in slash pine from half-sib tests were all strongly positive (Hodge and White 1992) within the range of ages investigated here, so this assumption seems logical.

Linear Model

The most complex model used was that for paired-site analyses of factorial tests, while all other models can be thought of as a subset of this model. This model was

$$y_{ijklmo} = \mu + t_i + b_{ij} + set_o + f_k + m_l + tf_{ik} + tm_{il} + fm_{kl} + tfm_{ikl} + p_{ijkl} + e_{ijklm}$$

where y_{ijklmo} is the m^{th} tree in the kl^{th} family in the o^{th} set, and j^{th} block of the i^{th} test,

μ is the population mean,

t_i is the random effect of the i^{th} test environment, $E(t_i)=0$ and $Var(t_i)=\sigma_t^2$,

b_{ij} is the random effect of the j^{th} block in the i^{th} test, $E(b_{ij})=0$ and $Var(b_{ij})=\sigma^2_b$,

set_o is the random effect of the o^{th} disconnected set of full-sib families, $E(set_o)=0$ and $Var(set_o)=\sigma^2_s$,

f_k is the random effect of the k^{th} female, $E(f_k)=0$ and $Var(f_k)=\sigma^2_{gca}$,

m_l is the random effect of the l^{th} male, $E(m_l)=0$ and $Var(m_l)=\sigma^2_{gca}$,

fm_{kl} is the random effect of the interaction between the k^{th} female and the l^{th} male, $E(fm_{kl})=0$ and $Var(fm_{kl})=\sigma^2_{sca}$,

tf_{ik} is the random interaction between the i^{th} test and k^{th} female, $E(tf_{ik})=0$, and $Var(tf_{ik})=\sigma^2_{tgca}$,

tm_{il} is the random interaction between the i^{th} test and l^{th} male, $E(tm_{il})=0$, and $Var(tm_{il})=\sigma^2_{tgca}$,

tfm_{ikl} is the random interaction between the i^{th} test and kl^{th} family, $E(tfm_{ikl})=0$, and $Var(tfm_{ikl})=\sigma^2_{tsca}$,

p_{ijkl} is the random effect of the $ijkl^{th}$ plot, $E(p_{ijkl})=0$ and $Var(p_{ijkl})=\sigma^2_p$, and

e_{ijklm} is the random effect within the $ijkl^{th}$ plot, $E(e_{ijklm})=0$ and $Var(e_{ijklm})=\sigma^2_w$.

In this model it was assumed that there was no covariance between the random effects in the model, that variances due to the female and male effects were equal, and that the female- and male-by-environment interactions were the same. GAREML produced one estimate of σ^2_{gca} and σ^2_{tgca} by pooling the estimates from the male and female parents. Note also, that where different aged test measurements were included in a paired-site analysis, all variance components defined above were viewed as covariance components.

In the case of diallel tests, individual parents are used as both males and females in the crosses. By assuming the absence of any reciprocal effects (i.e., that it does not matter whether

a parent is used as a male or as a female), it is possible to estimate the variance components for GCA and GCA-by-test location (Griffing 1956). Thus, when analyzing diallels, the terms f_k, m_l, tf_{ik} and tm_{il} in the above linear model were replaced by g_k, g_l, tg_{ik} and tg_{il} respectively for the general combining ability of the k^{th} (or l^{th}) parent, and fm_{kl} was replaced by s_{kl} for the specific combining ability. Finally, when conducting single-site analyses, the models used for factorial and diallel experiments were as described above except that all terms involving main and interaction effects due to the i^{th} test were dropped from the model.

Genetic Parameter Estimates

Four types of genetic parameters were estimated: heritability (h^2 and h_b^2 from pooled and single-site analyses, respectively), the proportion of dominance (d^2 and d_b^2), type B genetic correlation (r_B, which measures genotype-by-environment interaction [Burdon 1977]), and genetic correlation between ages (r_g). Type B and genetic correlations were estimated only from pooled analyses of site pairs.

From each pooled analysis of a pair of tests, both measured at the same age, heritability and the proportion of dominance were estimated as

$$h^2 = \frac{4\sigma^2_{gca}}{(2\sigma^2_{gca} + \sigma^2_{sca} + 2\sigma^2_{tgca} + \sigma^2_{tsca} + \sigma^2_{p} + \sigma^2_{w})}$$

$$d^2 = \frac{4\sigma^2_{sca}}{(2\sigma^2_{gca} + \sigma^2_{sca} + 2\sigma^2_{tgca} + \sigma^2_{tsca} + \sigma^2_{p} + \sigma^2_{w})}$$

In the absence of epistasis and maternal effects, with noninbred parents, these estimates of narrow sense heritability (h^2) and the proportion of dominance (d^2) are unbiased because i) σ^2_{gca}

is an estimate of one quarter of the additive genetic variance, σ^2_A (Cockerham 1963, Wright 1985, Cotterill *et al.* 1987), ii) σ^2_{sca} is an estimate of one quarter of the dominance genetic variance, σ^2_D (Cockerham 1963, Wright 1985, Cotterill *et al.* 1987), and iii) the denominator in both of the proceeding equations is an estimate of the total phenotypic variance, σ^2_P. Note that the ratio of additive to dominance variance (σ^2_A/σ^2_D) can be estimated by dividing h^2 by d^2. Details of the assumptions required in the genetic interpretation of the GCA and SCA variance components are described by Cockerham (1963) and Wright (1985).

When variance components are estimated from single-site analyses of progeny tests, it is impossible to separate the genetic (both additive and dominance genetic effects) from the genetic-by-environment interaction (GXE). Therefore estimates of variance due to general and specific combining abilities (from single sites) are biased upward (Comstock and Moll 1963) since the estimates include $\sigma^2_{gca} + \sigma^2_{tgca}$, and $\sigma^2_{sca} + \sigma^2_{tsca}$ respectively. Thus, for single-sites we can define the variance components attributable to the general and specific combining abilities as $\sigma^2_{GCA} = \sigma^2_{gca} + \sigma^2_{tgca}$, and $\sigma^2_{SCA} = \sigma^2_{sca} + \sigma^2_{tsca}$.

Upwardly biased heritability (h^2_b) and the proportion of dominance (d^2_b) were calculated separately for each measurement of all 171 progeny tests using the equations for paired-sites, above, but dropping all variance components involving interactions with tests and substituting σ^2_{GCA} and σ^2_{SCA} for σ^2_{gca} and σ^2_{sca}, respectively.

For all pairs of tests where both tests were measured at the same age, a Type B additive genetic correlation (Burdon 1977), termed r_B, was estimated in the following manner:

$$r_B = \frac{\sigma^2_{gca}}{(\sigma^2_{gca} + \sigma^2_{tgca})}.$$

This measures the degree of GXE at the additive level ($r_B = 1$ implies no GXE, while $r_B = 0$ implies that there is no consistent performance across test sites). Further, results from

unpublished data indicate that because the data were standardized, r_B is not influenced by scale effects in a similar manner to Dickerson (1962) and Yamada (1962) adjustments. Therefore r_B can be interpreted as GXE caused by parental rank changes across test sites.

From the results of the paired-test analyses it was also possible to estimate additive genetic correlations, in this case age-age genetic correlations, from each test-pair where both tests were measured at age_1 and age_2:

$$r_{g_{(age1,age2)}} = \frac{\sigma_{gca_{(age1,age2)}}}{\sqrt{(\sigma^2_{gca_{(age1)}}) \cdot (\sigma^2_{gca_{(age2)}})}}$$

where the additive genetic covariance between two ages ($\sigma_{gca(age1,\ age2)}$) is the quadratic average covariance for that test-pair, i.e.,

$$\sqrt{\sigma_{gca_{(age1,age2)}} \cdot \sigma_{gca_{(age2,age1)}}}$$

and $\sigma^2_{gca(age1)}$ and $\sigma^2_{gca(age2)}$ are estimated from the same test pair where data are the same age in both tests. This requires both tests to have been measured at both age_1 and at age_2.

Average Parameter Estimates

Average parameter estimates across all tests and test pairs, and an estimate of the standard error, need to be determined for biased and unbiased estimates of heritability, proportion of dominance, and genetic correlations. First, in order to prevent estimates of age-age correlation outside the theoretical limits from unduly influencing the means, estimates greater than 1.3 were set to 1.3 (Hodge and White 1992). Further, it was not necessary to distinguish between factorials and diallels, or between tests involving rust-free selections and those involving crosses between

the original first-generation selections, because for single sites there was no significant difference between these groups of tests for heritability or the proportion of dominance (both within age classes and across age classes).

Multiple regression was used to determine models to predict heritability and the proportion of dominance on the basis of site parameters such as the size of the test, mean rust infection level, age, site index, and coefficient of variation. However no model examined had an r^2 exceeding 0.2. Therefore, this approach was abandoned in favor of calculating average (unweighted mean) parameter estimates.

The use of unweighted means assumes that all estimates of a given parameter have equal variance. This, however, is unlikely to be true given the large test to test variation in the number of parents, families, blocks and trees/plot, and the different mating designs used. The use of weighted means is preferable under these circumstances, where each estimate is weighted by the inverse of its variance (as was done by Hodge and White 1992, and Woolaston *et al.* 1990). Using the asymptotic variances and covariances of the estimated variance components, it is possible to use a Taylor series expansion to approximate the variance of the genetic parameters described above (Kempthorne 1957; Becker 1975; Namkoong 1979; Hallauer and Miranda 1988). However, i) these approximate formulae provide no estimate of the variance when σ^2_{gca} (or σ^2_{GCA}) equals zero, ii) the Taylor series approximation of the variance of a genetic parameter was related to the size of the parameter estimate ($r=0.18$ for h^2_b, and $r=0.49$ for d^2_b), and iii) estimates were related to test size (refer chapter 2). These facts made the Taylor series approximates undesirable for use as weights. In the absence of suitable weights, unweighted means were used to calculate average parameters across all 171 tests. It is reasonable to expect that because of the large number of separate estimates of each genetic parameter, random fluctuations from estimate to estimate will tend to average to zero.

Results

Parameter Estimates

Averages of the biased estimates of heritability (h_b^2) for tree volume from single sites ranged from 0.117 at 5 years to a maximum of 0.155 at 11 years (Table 3-2). Estimates of the proportion of dominance (d_b^2) in these single-site analyses was relatively high when compared to h_b^2, ranging from 0.079 at 5 years to a maximum of 0.096 at 11 years, while the ratio of additive to dominance variance increased only slightly from 1.5 to 1.6 over the same period. In comparison to unbiased estimates of h^2 and d^2 obtained from paired-sites, all estimates from single sites were shown to be biased upwards. Unbiased heritability increased from 0.072 at 5 years to 0.12 at 11 years and older. Unbiased estimates of d^2 were proportionally smaller compared to h^2, than h_b^2 compared to d_b^2, and the ratio h^2/d^2 increased from 1.7 to over 3.1 between 5 and 14 years. Standard errors for all estimates were low: less than 0.02 in all but one case.

There is a moderate level of genotype-by-environment interaction present (Table 3-2). Type B genetic correlations for same-age test pairs changed from 0.61 at 5 years to 0.88 at 14 years, and estimates of biased heritability (h_b^2) exceeded unbiased estimates (h^2) by 1.5 times at 5 years, declining to 1.25 times at 11 years, and no difference at 14 years (Table 3-2). This indicates that although additive genetic-by-environment interaction was relatively large at 5 years of age, its importance declined with age and appeared to be relatively unimportant at 14 years (Type B not significantly different from 1.0). For estimates of the proportion of dominance, the biased estimates were approximately 1.7 times the unbiased estimates for all ages, i.e., dominance-by-environment interaction was approximately 70% of the size of the dominance genetic variance. This suggests that in this population, there is proportionately more dominance-by-environment interaction than additive-by-environment interaction (when compared to the

dominance and additive variances, respectively), and that the magnitude of the dominance-by-environment interaction does not decline as the trees age.

Table 3-2. Average heritability (h_b^2 and h^2) and the proportion of dominance (d_b^2 and d^2) for volume, as estimated from single-sites and pooled (paired-site) analyses, and Type B additive genetic correlation (r_B) estimated from full-sib slash pine progeny tests. Empirical standard errors are indicated in parentheses.

Age Class (Years)	Number of Tests		Heritability		Proportion of Dominance		h^2/d^2†		Type B (r_B)
	Single	Pooled	Single	Pooled	Single	Pooled	Single	Pooled	
5	142	372 (355)‡	0.117 (0.009)	0.072 (0.004)	0.079 (0.008)	0.044 (0.003)	1.48	1.66	0.609 (0.022)
8	118	368 (349)	0.142 (0.010)	0.100 (0.004)	0.084 (0.009)	0.051 (0.004)	1.68	1.94	0.677 (0.021)
11	85	222 (218)	0.155 (0.012)	0.124 (0.006)	0.096 (0.012)	0.055 (0.004)	1.61	2.25	0.761 (0.021)
14	21	15 (15)	0.123 (0.020)	0.122 (0.017)	0.066 (0.022)	0.039 (0.016)	1.85	3.14	0.882 (0.054)

† Equivalent to the ratio of additive to dominance variance (σ_A^2/σ_D^2)

‡ Number of test pairs used to calculate the mean Type B genetic correlation indicated in parentheses

Table 3-3. Average age-age genetic correlations (r_g) for volume, ± their empirical standard errors, estimated from full-sib tests pairs, from different age classes. (Estimates were bounded at 1.3, i.e., any individual estimate greater than 1.3 was arbitrarily set to 1.3.)

Age_1	Age_2	Number of Test Pairs	r_g (Age-Age)
5 Years	8 Years	205	0.885 ± 0.029
5 Years	11 Years	81	0.769 ± 0.045
5 Years	14 Years	11	0.565 ± 0.124
8 Years	11 Years	68	0.978 ± 0.031
8 Years	14 Years	9	0.737 ± 0.150
11 Years	14 Years	5	0.967 ± 0.063

Genetic correlations (r_g) between volume measurements at different ages are close to 1.0 when the difference in age is only 3 years, and declines to around 0.7-0.8 when there is a 6 year difference in age, and is approximately 0.6 if the difference in age is 9 years (Table 3-3).

Influence of Site Quality

The estimated site index at a base age of 25 years ranged from 44 ft to 101 ft (13.5 m to 31 m) in the 171 progeny tests. Given this large range in the quality of the sites, it might be expected that site quality had some effect on estimated genetic parameters (h^2, d^2 and r_B). However, an analysis of variance of the single-site heritability (h_b^2) revealed that although differences in the age classes are significant, there were no significant differences between the site index classes, and the interaction between age classes and site index classes also was not significant (using SAS Type III sums of squares). For the biased estimates of the proportion of dominance (d_b^2) from single-site analyses neither age class, site index class nor the interaction was significant.

For biased heritability, although the differences between the three site index classes was not significant, there appears to be a general increase in the heritability as site index increases. The mean biased heritability in the low, medium, and high site index classes (across all age classes) were 0.108 ± 0.011, 0.136 ± 0.008, and 0.158 ± 0.010, respectively. Thus h_b^2 seems to increase 0.02 to 0.03 between low to medium and medium to high site index classes. A simple linear regression model between h_b^2 and site index classes had a low r^2 (0.026) but was highly significant (p-value = 0.002 for the F test). This indicates that although there is a lot of variation not explained by the site index classes, the average effect is probably still important, and suggests that the degree of genetic control over the expression of volume may increase as site index increases. Estimates of d_b^2 show no relationship to site index.

To further investigate the nature of GXE, pairs of connected tests were classified as having the 'same' site index if the difference in site index was less than 8.5 ft (2.6 m), otherwise as having 'different' site indices. A difference in site index of 8.5 ft was chosen to ensure comparability with previous results from open-pollinated tests reported by Hodge and White

(1992). Analysis of variance of the Type B genetic correlation (r_B) revealed no significant interaction between the age classes and the 'same'/'different' grouping. Also the difference in r_B between test pairs with the 'same' site index was not significantly different from test pairs of 'different' site index. However the differences in r_B between age classes (Table 3-2) was highly significant (p-value = 0.0001). This contrasts with results reported by Hodge and White (1992) who found for all growth traits, test pairs of similar site index had higher average r_B than test pairs of dissimilar site index.

Influence of the Level of Rust Infection and Survival

The fraction of trees infected with fusiform rust was calculated for each test. Tests were then allocated to one of three classes: less than 35% rust, 35-65% rust, and greater than 65% rust infection. For both biased heritability and the biased proportion of dominance, the interaction between age and rust classes was not significant. But only for d_b^2 was the effect of the rust class found to be significant (p-value = 0.0005) in an analysis of variance. The mean d_b^2 across all age classes were: 0.074 ± 0.007, 0.080 ± 0.008, and 0.142 ± 0.020 for the low, medium and high rust classes respectively. The low and medium rust classes were not significantly different from one another; however, the high rust class (>65% rust) was significantly higher than the other two (based on Duncan's multiple range test).

The increase in the proportion of dominance at high rust levels was observed most strongly at 8 and 11 years, and corresponds with mean survival (over all tests) which was less than 65 percent. It is therefore possibly only a consequence of the poor survival in these tests which results from high levels of rust infection, rather than a consequence of the rust itself. However if all tests with less than 65 percent survival are eliminated, and the analysis of variance of d_b^2 is repeated, the results are the same as when all the tests were used. The average d_b^2 is still

significantly greater in the high rust tests, than both lower rust classes. The mean survival in the three rust classes in this second analysis were: 86.2%, 80.7% and 76.9% in the low, medium and high rust classes respectively. High levels of rust infection, therefore, do appear to inflate estimates of dominance variance for volume in slash pine.

Discussion

Heritability Estimates

The average biased and unbiased heritability estimates obtained for volume from the full-sib tests are similar to previous estimates from the CFGRP open-pollinated tests (Hodge and White 1992) and within the range of heritabilities reported for growth traits in slash pine and other *Pinus* species (Matziris and Zobel 1973, Dorman and Squillace 1974, Snyder and Namkoong 1978, Bridgwater *et al.* 1983, Lambeth *et al.* 1983, Foster 1986, Sluder 1986, Talbert 1984, Tauer and McNew 1985, Dean *et al.* 1986, Cotterill *et al.* 1987, Cotterill and Dean 1988, Riemenschneider 1988, Falkenhagen 1989, Balocchi 1990, Johnson and Burdon 1990, Woolaston *et al.* 1990, Ladrach and Lambeth 1991, Barnes *et al.* 1992, Hodge and White 1992, Balocchi *et al.* 1993, Pswarayi 1993; see Appendix), and forest trees in general (Cornelius 1994). It is difficult to make comparisons between the present study and most published heritability estimates because estimates in the literature are frequently obtained from only one or two tests and rarely from more than five tests, while the current study includes data from 171 tests. Individual heritability estimates obtained from this study range between 0 and 0.36 (for h^2) or 0.46 (for h_b^2). Thus, if only one test, or a small group of tests were considered, then almost any estimate of heritability (within the limits 0.0 to 0.5) could be possible. Hence comparisons will be

restricted to estimates published by Hodge and White (1992), who used data from 57 open-pollinated tests to estimate average genetic parameters in slash pine.

The heritability estimates (especially the single-site biased estimates) are lower than Hodge and White's estimates by 0.02 to 0.05. There are a number of possible reasons for this observed difference between the estimates from the open-pollinated and control-pollinated tests, including: i) the open-pollinated progenies may not be all half-sibs; ii) differences in the estimation procedure; and, iii) possible mistakes when making the full-sib families.

Some authors argue that it may have be more appropriate to multiply the variance component for open-pollinated families by three, rather than four, to obtain an estimate of the additive variance (Squillace 1974, Adams and Joyce 1990, Vargas-Hernandez and Adams 1991). Such as adjustment would reduce the heritability estimates from the open-pollinated tests by 25 percent, e.g., adjusted unbiased heritability estimates would be 0.060, 0.117 and 0.121 at five, ten and fifteen years. Such adjusted estimates are very close to estimates obtained from the full-sib tests. This might be coincidental as there are many other aspects of open-pollinated families that could have opposite effects on heritability estimates (Sorensen and White 1988).

During the estimation procedure Hodge and White (1992) eliminated approximately 10 percent of tests because the p-value (of the F-test for the family effect in the analysis of variance) was less than 0.25, thereby eliminating tests with very low heritability from the analysis, and so slightly increasing heritability estimates. However the effect of this is only small. Byram and Lowe (1986) in their study of 30 full-sib loblolly pine (*P. taeda*) tests, similarly excluded tests where the p-value for crosses in the analysis of variance was less than 0.10, or if there were less than two crosses per parent. In the current study no analysis of variance was conducted, but if 10 percent of tests with the lowest heritability (h_b^2) or tests with less than two crosses per parent are eliminated, then the average heritabilities change only very slightly.

A third possible reason for the differences between heritability estimates from full-sib tests and open-pollinated tests may have been errors in making the control-pollinated families. Pollen contamination may result if female strobili are bagged too late. Mistakes can occur in the records and labels from pollination, to seed collection and sowing in the nursery, through to eventual planting in the field. Such mistakes do occur in operational tree improvement programs (Adams *et al.* 1988) and may go undetected; however, it is impossible to accurately quantify their possible impact on the heritability estimates obtained in this study.

Proportion of Dominance

There are relatively few genetic studies of the importance of dominance variance in forest trees, and even fewer with estimates of dominance at more than one age. Some of the most comprehensive studies of the importance of dominance in forest trees at multiple ages include: Byram and Lowe (1986), Balocchi (1990) and Balocchi *et al.* (1993) in loblolly pine; Pswarayi (1993) in slash pine; and Samuel (1991) in sitka spruce (*Picea sitchensis*). In these studies greatly divergent patterns in the ratio of additive to dominance variance (σ_A^2/σ_D^2) are evident as the trees age (Figure 3-1). From these studies no general pattern in the ratio σ_A^2/σ_D^2 is evident across species, or even within a species. Except for the study by Byram and Lowe (1986) which used relatively few tests after age 5, all other studies reveal a predictable pattern in the ratio, either upwards or downwards with increasing age. This trend is generally linear except for Balocchi (1990), where a curvilinear pattern is evident at ages less than five years.

The two studies of dominance in slash pine, Pswarayi (1993) and the current study, have markedly different results. Pswarayi (1993) found that the amount of additive variance declined with age when compared to the amount of dominance variance, thus dominance becomes increasingly important at later ages. This is the exact opposite to what was found in the current

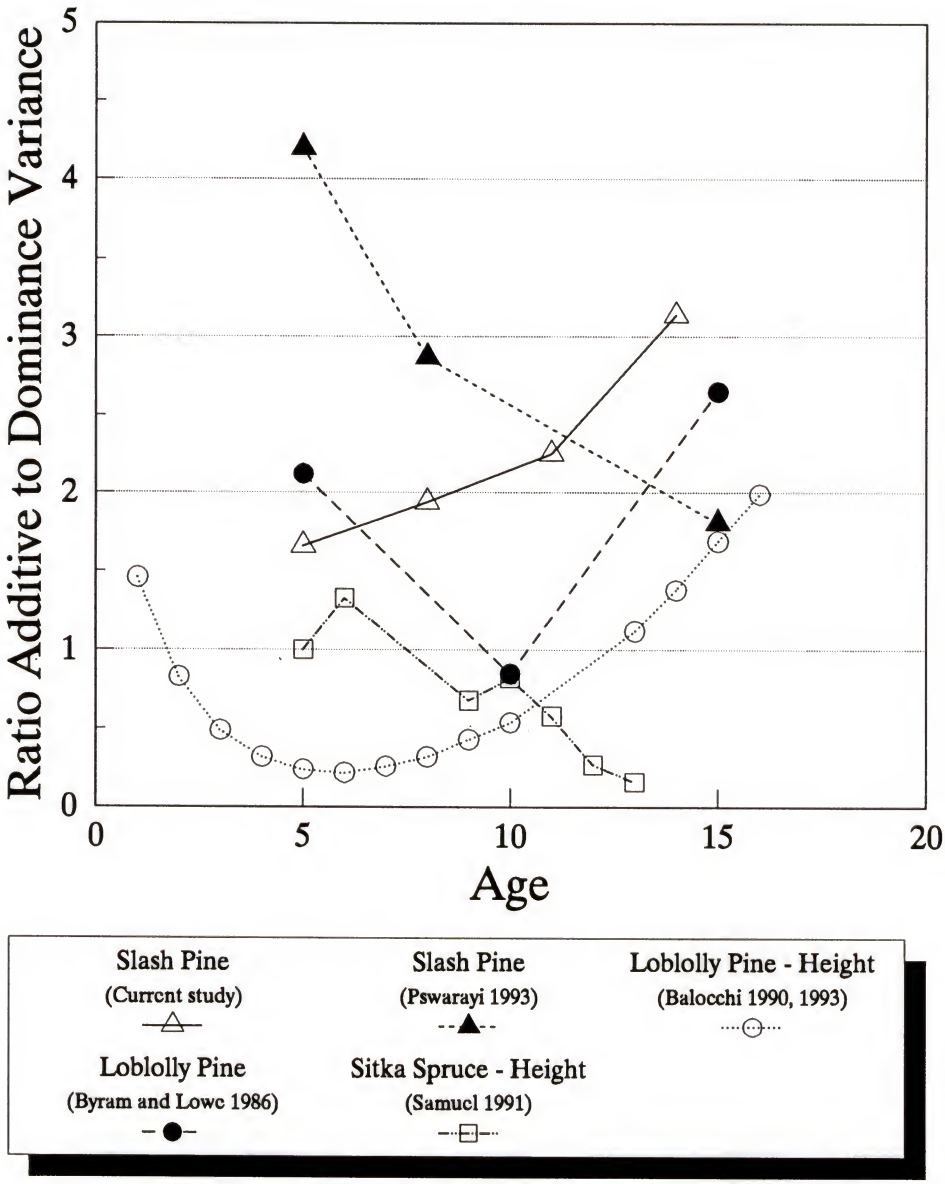

Figure 3-1. Estimates of the ratio of additive to dominance variance in full-sib tests of slash pine (current study and Pswarayi 1993), loblolly pine (Byram and Lowe 1986, Balocchi 1990, Balocchi *et al.* 1993), and sitka spruce (Samuel 1991).

study, where dominance was found to decline in importance relative to additive variance with increasing age. Possible explanations for this marked difference could be: the small number of tests involved in Pswarayi's study (only three tests, but each test quite large); and the fact that these two studies represent data from two distinct populations of slash pine, growing on two different continents.

Age-Age Correlations

The model developed by Lambeth (1980) to predict age-age height correlations in the *Pinaceae* was applied to the age-age genetic correlations (r_g) for volume obtained in this study (Figure 3-2). Lambeth's model is based on the simple linear regression between the natural logarithm of the age ratio (young age divided by older age, termed LAR) and the age-age phenotypic correlation of height growth across a number of species in the *Pinaceae*.

The predictive model for age-age volume genetic correlations in slash pine from open-pollinated data is very similar to that obtained from the current study (Figure 3-2): the slopes are almost identical and the intercept terms differ by only 0.1. The absolute values of the estimates obtained in these two studies are also very similar, and both studies differ substantially from Lambeth's model. Nevertheless, in all cases the LAR proved to be a good predictor of the age-age correlations.

Determination of the optimum selection age is not a simple matter; however, examination of the age-age correlations and the heritability estimates from this study tend to suggest: i) if the target trait is 14 year volume, then there is no loss in selection efficiency by selecting at 11 years because the heritabilities are equal and the genetic correlation is almost equal to 1.0; and ii) age-age correlations between 8 year and 14 year volume are relatively strong, and there is little loss in heritability. Using formulae for indirect selection (Falconer 1989, p.320), the parameter

estimates in Tables 3-2 and 3-3, and assuming the same selection intensity at the two ages, the efficiencies of early selection compared to selecting at 14 years are 0.33, 0.60 and 0.98, for selection at 5, 8 and 11 years. From these data, it would therefore appear that there is no value in delaying selection beyond 11 years, and depending on the selection method and the economic assumptions even earlier selection may be warranted.

GXE and Influence of Site Quality

The results presented indicate the existence of moderate amounts of genotype-by-environment interaction; more specifically additive genetic-by-environment interaction. The importance of this interaction is related to age but no significant relationship could be found with site quality. Between 5 and 14 years the estimated Type B genetic correlation increases from 0.609 to 0.882 (Table 3-2). This contrasts with previous results from open-pollinated tests of slash pine. Hodge and White (1992) found no increase in the Type B genetic correlation between 5 and 15 years (estimated Type B genetic correlations for volume were 0.592, 0.665 and 0.628, respectively at 5, 10 and 15 years in slash pine); however, they found site pairs of similar site index to have consistently higher r_B than test pairs with different site indices. No reasons for these differences between these two studies are apparent.

As a rule of thumb, Shelbourne (1972) suggested that when the variance of GXE is at least half the size of the additive variance (i.e., Type B = 0.67), breeding strategies should be modified to address this interaction, or potential genetic gains will be compromised. Estimates from this study for 5 and 8 year data are close to this lower limit; however, later age estimates of the Type B genetic correlation for volume are considerably higher. It is therefore possible that the existence of GXE at young ages may be of no practical concern, and so can be ignored. This

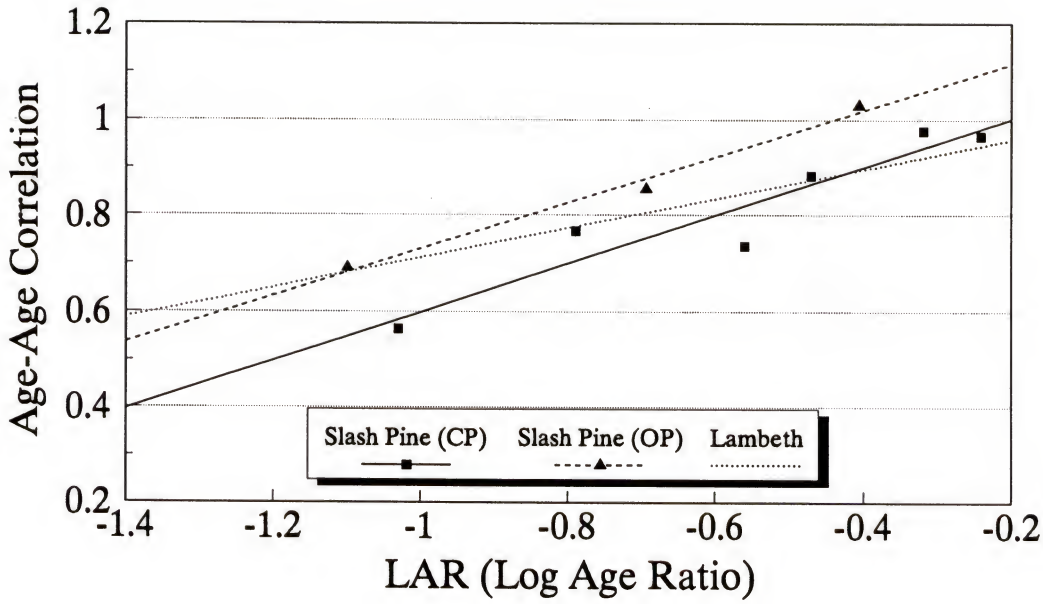

Figure 3-2. Age-age genetic correlation (r_g) in slash pine predicted as a simple linear function of the natural logarithm of the age ratio (LAR), as developed by Lambeth (1980). The regression models are: i) the current study of slash pine control-pollinated (CP) slash pine, $\hat{r}_g = 1.212 + 0.482(LAR)$, $r^2 = 0.894$, using volume; ii) open-pollinated (OP) slash pine (Hodge and White 1992), $\hat{r}_g = 1.212 + 0.482(LAR)$, $r^2 = 0.987$, using volume; and, iii) Lambeth (1980), $\hat{r} = 0.102 + 0.308(LAR)$, using height.

however needs to be verified through the collection of older data from these full-sib tests: currently the bulk of the data is from 5- and 8-year-old tests.

There was some indication that site quality may influence the estimated biased heritability (h_b^2), with sites of higher site index generally giving higher estimates of the heritability. This observation also contrasts with results reported by Hodge and White (1992) who found no relationship between site index and h_b^2; however, the present study includes almost 120 more tests. If better quality sites do in fact lead to higher heritability estimates, this would suggest that more genetic gain will result from establishing progeny tests on sites with high site index. However such a testing strategy also requires low genotype-by-environment interaction. The Type B genetic correlations estimated for test pairs with the 'same' and 'different' site index revealed that for 5- and 8-year-old tests there is no more GXE between sites with the same site index and those of different site index. Also at later ages, there is considerably less GXE between the tests of similar site index, and the higher heritability associated with sites of better quality may offset any loss of gain due to GXE. Further investigation is required to determine the optimum testing strategy, since in general the presence of GXE argues against testing only on good quality sites.

These results also indicate the existence of dominance genetic-by-environment interaction. The biased estimates of the proportion of dominance are always higher than the unbiased estimates (Table 3-2), with the ratio of biased to unbiased estimates in the range 1.6 to 1.8. In contrast to the additive-by-environment interaction, the dominance interaction does not appear to decline with importance with age. Therefore, deployment options involving the use of full-sib families used to capture gain resulting from dominance will most probably need to pay more careful attention to site differences than conventional strategies which only aim to capture gain from additive effects.

Implications for the Breeding Strategies

The amount of dominance variance in volume of slash pine is one half to one third that of additive variance. In a survey of loblolly pine experiments McKeand *et al.* (1986) similarly found that dominance variance was an important part of the total genetic variation in traits associated with yield. Other evidence for the existence of significant amounts of dominance in the CFGRP slash pine population is found in results from inbreeding studies, where for every 0.1 increase in the inbreeding coefficient (Falconer 1989, p.62) the inbreeding depression for volume was between 6 and 19 percent (Hodge *et al.* 1994). Theoretically, inbreeding depression occurs in the presence of directional dominance (Falconer 1989, p.251).

The current breeding strategy in the CFGRP slash pine breeding program concentrates on recurrent selection for GCA, but does not allow for making genetic gains from dominance in the main breeding population (White *et al.* 1993). This decision was based on the assumption that dominance was less important than additive variance and information from loblolly pine that dominance becomes less important at later ages. These assumptions have been validated by the current study, up to age 14. When tests with poor survival are eliminated there is some evidence to suggest that the ratio of additive to dominance variance stabilizes after 8-11 years, rather than continuing to increase, and may not increase above 3.0.

Under the current strategy, the only full-sib families that will be tested in replicated experiments are a total of 30 crosses produced by single pair mating 30 individuals in one elite population with 30 individuals in the other elite population (White *et al.* 1993). The data do not indicate that any great deviation from this strategy is necessary, but given that dominance is present, comprising about 30 percent of the total genetic (additive plus dominance) variation in volume at 8 to 11 years, a larger investment in full-sib testing may be warranted. Individual crosses that combine both good parental GCA and large positive SCA will be relatively rare;

therefore, testing of more families may be necessary. There would appear to be little profit in progeny testing crosses between parents which are not in the elite populations because the chance is small that additional gain from positive SCA in a given cross will be sufficient to overcome the lower general combining effects of parents outside of the elite populations.

A form of reciprocal recurrent selection to breed for SCA as well as GCA has been proposed for loblolly pine using paired sublines (McKeand *et al.* 1986). A strategy similar to this could be applied to the two elite populations of the CFGRP slash pine improvement program by testing each individual in one elite population with a bulked pollen mix from the other elite population (and *vice versa*). The parents in the elite populations were to be GCA tested under the existing strategy; therefore, the only additional work would be forming the two pollen mixes of the parents in the two elite populations. The parents in the elite populations are at the upper extreme of the distribution of breeding values; therefore, it is possible that the average estimates of dominance variance obtained in this study from the entire first-generation breeding population may not accurately reflect the situation in the elite populations. The actual levels of dominance could more or less than estimated here. The reciprocal testing will select for parents that combine well with the other elite population, and if dominance is important divergent gene frequencies can be expected to develop; however, if GCA is predominant then individuals of high GCA will be selected (McKeand *et al.* 1986), and so nothing is lost.

Additional gains, therefore, could be realized through the use of i) more intensive testing of full-sib crosses between the two elite populations, and/or ii) the use of a modified reciprocal recurrent selection scheme between the two elite populations. However further work is required to determine whether the cost of the additional work is likely to be offset by any additional gain. Also unless deployment of either full-sib families or clones becomes operationally feasible with slash pine in the southeastern United States, any gains from dominance will be of no value.

However, given that a suitable propagation system is developed within the next ten years, even small additional gains will have substantial economic value because of the large slash pine plantation estate owned by members of the CFGRP.

Other results from this study generally confirm previous results from the open-pollinated tests: estimates of heritability, genotype-by-environment interactions, age-age correlations are similar in the two types of tests. Therefore this study provides no evidence to modify the breeding strategy (White *et al.* 1993) other in the areas relating to the capture of gains from dominance variance previously discussed.

Conclusions

Heritability estimates for volume in slash pine are relatively low compared to published estimates: narrow sense heritability at five years is 0.07 and increases to around 0.12 at 11 and 14 years of age. These estimates are very similar to previous estimates from open-pollinated tests reinforcing the use of such testing methods to estimate additive variance and heritability. The amount of dominance variance found in this population of slash pine was relatively large with around 1.6 times as much additive variance as dominance variance at 5 years, increasing to over two times at 11 years of age or older. The existence of dominance variance may mean that some changes in the breeding strategy are necessary in order to capture additional gain from this dominance variance; however, more work is required to determine the economic value of breeding and testing for SCA. The levels of genotype-by-environment interaction are similar to previous estimates and were found to have some relationship to site quality. Estimates of Type B genetic correlations increased from around 0.6 at 5 years to over 0.8 at 14 years of age, and so the importance of genotype-by-environment interaction appears to decline with age. Further studies are required to confirm this as relatively few tests in this study were measured at 11 and

14 years of age. The levels of GXE found support the testing of progeny on many sites as prescribed under the current strategy. Age-age correlations were likewise found to be similar to results from the open-pollinated tests and could be predicted as a linear function of the natural logarithm of the age ratio as in other pines.

CHAPTER 4
GENETIC PARAMETER ESTIMATES FOR RESISTANCE TO RUST (*CRONARTIUM QUERCUUM*) INFECTION FROM FULL-SIB TESTS OF SLASH PINE (*PINUS ELLIOTTII*), MODELLED AS A FUNCTION OF RUST INCIDENCE

Introduction

The most serious disease affecting slash (*Pinus elliottii* Engelm. var *elliottii*) and loblolly pines (*P. taeda* L.) within their natural ranges is fusiform rust (Walkinshaw and Anderson 1988, Walkinshaw and Roland 1990). Fusiform rust is caused by the fungus *Cronartium quercuum* (Berk.) Miyabe ex Shirai f. sp. *fusiforme*, and infects many of the southern pines. Of the commercially important *Pinus* species grown in the southern United States, slash pine is one of the most susceptible to fusiform rust (Schmidt *et al.* 1981), and the economic impacts of the disease are considerable. Yield losses result from mortality and stem breakage (following the formation of stem-girdling galls), reduced marketability, and decreased growth rates (Schmidt *et al.* 1981, Hodge and White 1986, Walkinshaw and Roland 1990). Annual losses in commercial slash pine and loblolly pine (*P. taeda* L.) stands have been estimated at 562 million board feet (1.3 million cubic meters) of sawtimber and 194 million cubic feet (5.5 million cubic meters) of growing stock (Phelps and Czabator 1980), and have been estimated to cost approximately 9 and 26 million dollars annually for slash and loblolly pine, respectively, grown in Florida, Georgia, South Carolina, North Carolina and Virginia (Anderson *et al.* 1986).

Although it is known that useful levels of genetic variation in resistance to fusiform rust infection exist within slash pine (Dinus and Griggs 1975, Goddard *et al.* 1975, Schmidt *et al.* 1981, Walkinshaw and Bey 1981), and substantial genetic gains have been achieved in the resistance of slash pine to fusiform rust (Sluder 1986, Hodge *et al.* 1989, Hodge *et al.* 1990), relatively little concerning the quantitative genetics of rust resistance in slash pine has been published. Rockwood and Goddard (1973) obtained individual tree heritability estimates ranging from 0.035 to 0.262 in ten progeny tests with an average of 0.167, Sohn and Goddard (1979) report heritabilities in the range 0.1 to 0.4 in eight tests with an average of 0.25, and Hodge *et al.* (1990) report realized heritabilities between 0.3 and 0.4. And finally, in two factorial tests Kraus (1973) found a significant amount of dominance variance in the number of rust galls observed per tree.

In other *Pinus* species there is likewise relatively little published information on the quantitative genetics of disease resistance. Sluder (1988, 1993) estimated the family heritability of fusiform rust resistance in a single loblolly pine test to be 0.69 at both 10 and 15 years of age, and found that the amount of dominance variance was relatively small compared to additive variance. In nine progeny tests of *P. radiata* the individual tree heritability of resistance to *Dothistroma* needle blight varied between 0.13 and 0.40 (average of 0.26), and dominance was very small compared to additive variance (Carson 1989). And in a single test of *P. muricata*, individual tree heritability of resistance to *Dothistroma* needle blight was estimated to be 0.29 (Ades *et al.* 1992).

There is evidence to suggest that genetic parameter estimates are influenced by the mean infection level, as might be expected given the binomial nature of rust infection (rust infection is generally scored on the basis of presence = 1 and absence = 0). When rust incidence is near either extreme (0% or 100% infection) there is little genetic or phenotypic variation, and

consequently genetic parameters will be poorly estimated. Some authors have reported a positive association between the heritability of rust resistance in slash pine and the mean level of rust infection (Rockwood and Goddard 1973, Sohn and Goddard 1979, White and Hodge 1989, p.185). Genotype-by-environment interaction of rust resistance is not thought to be important in slash pine, even though statistically significant in some cases (Goddard and Schmidt 1979). Nevertheless family mean correlations between two wind pollinated tests with reasonable levels of rust infection (over 20%) are higher when both tests have similar levels of rust infection (Schmidt and Goddard 1971, Sohn *et al.* 1975). Therefore, it appears that genetic parameter estimates must be considered in conjunction with the abundance and virulence of the pathogen as measured by the average infection level in the stand.

This paper used data from 171 separate full-sib tests of slash pine established by the Cooperative Forest Genetic Research Program (CFGRP) based at the University of Florida to i) provide reliable genetic parameters estimates for rust resistance in slash pine and ii) investigate the relationship between the mean incidence of rust in a test and the genetic parameter estimates. The genetic parameters included in this study are biased (single-site) and unbiased (paired-site) estimates of heritability, the proportion of dominance, type B genetic correlations (a measure of GXE [Burdon 1977]), and age-age genetic correlations.

Materials and Methods

Full-sib Progeny Test Data

First-generation slash pine parents used by the CFGRP in full-sib matings originated from two separate phases of mass selection: the first (1956-1963) concentrated on stands with a low incidence of fusiform rust, while the second (early 1970s) was restricted to stands where the average fusiform rust infection exceeded 70% (Hendrickson 1976). Some 2500 and 550

phenotypically superior trees were selected in the first and second phases, respectively. The second phase of mass selection, of 'rust-free' trees in high rust-incidence areas, was necessary because relatively few of the initial first-generation parents proved resistant to fusiform rust (Goddard 1980). The CFGRP established over 200 replicated full-sib progeny tests between 1966 and 1989, which involve i) factorial matings amongst the initial first-generation trees, ii) diallel or factorial matings among only the better quality initial first-generation trees, and iii) factorial crosses using the rust-free trees as male parents and the better initial first-generation trees as female parents.

As of April 1994, 171 different full-sib tests were old enough (more than three years old) to be included in this study. These tests are mainly in Florida, Georgia, and Alabama, but a few tests occur in Mississippi and South Carolina. All tests were established in a randomized complete block design with 3-12 blocks (mean=5.4 blocks), and each full-sib family was represented by one plot of 5-10 trees (mean=6.9 trees) arranged as either a row-plot or a noncontiguous plot. In any one test there were between 6 and 86 full-sib families (mean=30.6 families), which were derived from crosses amongst 6 to 47 parents (mean=18.2 parents). In total the tests included progeny from over 700 parents (165 rust-free parents), represented by over 2100 different full-sib families (550 families involving rust-free parents), and approximately 170,000 individual trees.

The presence (score=100) or absence (score=0) of rust infection was assessed visually on each tree at approximately three year intervals between 4 and 15 years of age, with individual tests often assessed at least twice. Data were grouped into one of four, 3-year age classes, centered at 5, 8, 11 and 14 years, and each of the 362 test-age class combinations were treated as separate data points. Tests were also grouped into five classes based on the mean level of rust infection in the test (0-20%, 20-40%, ...).

Estimation of Variance Components

The presence of considerable imbalance within this set of full-sib data resulting from incomplete mating designs, unequal representation of parents in tests, and mortality precluded the use of analysis of variance based methods to estimate variance components. Restricted maximum likelihood (REML) estimation (Patterson and Thompson 1971) is generally considered to be a better choice in such circumstances (Searle *et al.* 1992). GAREML, a computer program developed by Huber (1993) which utilizes Giesbrecht's algorithm (1983) to provide REML estimates, has been shown to provide variance component estimates with desirable properties when applied to data sets with the amount and type of imbalance commonly found in genetic tests of forest trees (Huber 1993, p.82). Therefore GAREML was used to analyze the individual tree rust scores from these 171 full-sib slash pine tests.

REML variance components estimates were obtained from i) single-site analyses, and ii) analysis of connected pairs of tests (i.e., test pairs with at least five common parents). All possible connected test pairs were analyzed, utilizing data from all ages. However data of different ages obtained from the same test were never analyzed together, thereby avoiding the potential problem of nonindependent error terms due to environmental correlation within a single site (Hodge and White 1992). When test pairs were constructed using data from two tests measured at the same age, the REML estimates are variance components; however, when the pair of tests are of different age the estimates obtained are considered to be covariance components. This method of analysis forces covariance components to be nonnegative, in the same way that (REML) variance components are nonnegative.

For binomial traits, the unit of analysis can be family means (possibly combined with a transformation such as arcsin or logistic [Sohn and Goddard 1979, De Souza 1990, De Souza *et al.* 1991]) because such means are approximately normally distributed by the central limit

theorem (Mendenhall *et al.* 1981, p.325). However, as one of the ultimate goals of this work was the prediction of genetic gain from within family selection, estimates of the individual heritability and the within plot variance were required. Further, i) the approximate variance of heritability estimates, which depends on underlying normality assumptions, was found to be closely related to empirical estimates of variance even when individual binomial rust scores were used in the analysis (refer chapter 2), ii) results from previous studies have demonstrated that using an arcsin transformation of the family means does not greatly affect the relative size of variance component estimates (Rockwood and Goddard 1973, Sohn and Goddard 1979), iii) REML estimation has been shown to be relatively robust to violations of the underlying normality assumptions (Banks *et al.* 1985; Westfall 1987), iv) Banks *et al.* (1985) using categorical data (with two to six classes) generated from an underlying normal distribution, demonstrated that REML estimation was acceptable, at least in terms of heritability estimates, and v) in simulation studies using REML estimation, individual observations have been shown to be generally superior to the use of plot means as the unit of analysis (Huber 1993). Therefore, for these reasons and to avoid the problems of back-transformation, all analyses were conducted using untransformed individual rust scores.

Linear Model

The most complex model used was that for paired-site analyses of factorial tests, while all other models can be thought of as a subset of this model. This model was

$$y_{ijklmo} = \mu + t_i + b_{ij} + set_o + f_k + m_l + tf_{ik} + tm_{il} + fm_{kl} + tfm_{ikl} + p_{ijkl} + e_{ijklm}$$

where y_{ijklmo} is the m^{th} tree in the kl^{th} family in the o^{th} set and j^{th} block of the i^{th} test,

μ is the population mean,

t_i is the random effect of the i^{th} test environment, $E(t_i)=0$ and $Var(t_i)=\sigma^2_t$,

b_{ij} is the random effect of the j^{th} block in the i^{th} test, $E(b_{ij})=0$ and $Var(b_{ij})=\sigma^2_b$,

set_o is the random effect of the o^{th} disconnected set of full-sib families, $E(set_o)=0$ and $Var(set_o)=\sigma^2_s$,

f_k is the random effect of the k^{th} female, $E(f_k)=0$ and $Var(f_k)=\sigma^2_{gca}$,

m_l is the random effect of the l^{th} male, $E(m_l)=0$ and $Var(m_l)=\sigma^2_{gca}$,

fm_{kl} is the random effect of the interaction between the k^{th} female and the l^{th} male, $E(fm_{kl})=0$ and $Var(fm_{kl})=\sigma^2_{sca}$,

tf_{ik} is the random interaction between the i^{th} test and k^{th} female, $E(tf_{ik})=0$, and $Var(tf_{ik})=\sigma^2_{tgca}$,

tm_{il} is the random interaction between the i^{th} test and l^{th} male, $E(tm_{il})=0$, and $Var(tm_{il})=\sigma^2_{tgca}$,

tfm_{ikl} is the random interaction between the i^{th} test and kl^{th} family, $E(tfm_{ikl})=0$, and $Var(tfm_{ikl})=\sigma^2_{tsca}$,

p_{ijkl} is the random effect of the $ijkl^{th}$ plot, $E(p_{ijkl})=0$ and $Var(p_{ijkl})=\sigma^2_p$, and

e_{ijklm} is the random effect within the $ijkl^{th}$ plot, $E(e_{ijklm})=0$ and $Var(e_{ijklm})=\sigma^2_w$.

In this model it was assumed that there was no covariance between the random effects, that variances due to the female and male effects were equal, and that the female- and male-by-environment interactions were the same. GAREML produced one estimate of σ^2_{gca} and σ^2_{tgca} by pooling the estimates from the male and female parents. Note also, that where different aged test measurements are included in a paired-site analysis, all variance components defined above are viewed as covariance components.

In the case of diallel tests, individual parents are used as both males and females in the crosses. By assuming the absence of any reciprocal effects (i.e., that it does not matter whether

a parent is used as a male or as a female), it is possible to estimate the variance components for GCA and GCA-by-test location (Griffing 1956). Thus, when analyzing diallels the terms f_k, m_l, tf_{ik} and tm_{il} in the above linear model were replaced by g_k, g_l, tg_{ik} and tg_{il} respectively for the general combining ability of the k^{th} (or l^{th}) parent, and fm_{kl} was replaced by s_{kl} for the specific combining ability. Finally, when conducting single-site analyses, the models used for factorial and diallel experiments were as described above, except that all terms involving main and interaction effects due to the i^{th} test were dropped from the model.

Genetic Parameter Estimates

Four types of genetic parameters were estimated: heritability (h^2 and h_s^2 from pooled and single-site analyses, respectively), the proportion of dominance (d^2 and d_s^2), type B genetic correlation (r_B), which measures genotype-by-environment interaction, and genetic correlation between ages (r_g). Type B and genetic correlations were estimated only from pooled analyses of site pairs.

From each pooled analysis of a pair of tests, both measured at the same age, heritability and the proportion of dominance were estimated as

$$h^2 = \frac{4\sigma^2_{gca}}{(2\sigma^2_{gca} + \sigma^2_{sca} + 2\sigma^2_{tgca} + \sigma^2_{tsca} + \sigma^2_{p} + \sigma^2_{w})}$$

$$d^2 = \frac{4\sigma^2_{sca}}{(2\sigma^2_{gca} + \sigma^2_{sca} + 2\sigma^2_{tgca} + \sigma^2_{tsca} + \sigma^2_{p} + \sigma^2_{w})}$$

In the absence of epistasis and maternal effects, with noninbred parents, these estimates of narrow sense heritability (h^2) and the proportion of dominance (d^2) are unbiased, because i)

σ^2_{gca} is an estimate of one quarter of the additive genetic variance, σ^2_A (Cockerham 1963, Wright 1985, Cotterill *et al.* 1987), ii) σ^2_{sca} is an estimate of one quarter of the dominance genetic variance, σ^2_D (Cockerham 1963, Wright 1985, Cotterill *et al.* 1987), and iii) the denominator in both of the proceeding equations is an estimate of the total phenotypic variance, σ^2_P. Note that the ratio of additive to dominance variance (σ^2_A/σ^2_D) can be estimated by dividing h^2 by d^2. Details of other assumptions required in the genetic interpretation of the GCA and SCA variance components are described by Cockerham (1963) and Wright (1985).

When variance components are estimated from single-site analyses of progeny tests, it is impossible to separate the genetic (both additive and dominance genetic effects) from the genetic-by-environment interaction. Therefore estimates of variance due to general and specific combining abilities (from single-sites) are biased upward (Comstock and Moll 1963), since the estimates include $\sigma^2_{gca} + \sigma^2_{tgca}$, and $\sigma^2_{sca} + \sigma^2_{tsca}$ respectively. Thus for single-sites we can define the variance components attributable to the general and specific combining abilities as: $\sigma^2_{GCA} = \sigma^2_{gca} + \sigma^2_{tgca}$, and $\sigma^2_{SCA} = \sigma^2_{sca} + \sigma^2_{tsca}$.

Thus upwardly biased heritability (h^2_b) and the proportion of dominance (d^2_b) were calculated separately for each measurement of all 171 progeny tests using the equations for paired-sites given above, but dropping all variance components involving interactions with tests, and substituting σ^2_{GCA} and σ^2_{SCA} for σ^2_{gca} and σ^2_{sca} respectively.

For all pairs of tests where both tests were measured at the same age, a Type B additive genetic correlation (Burdon 1977), termed r_B, was estimated in the following manner:

$$r_B = \frac{\sigma^2_{gca}}{(\sigma^2_{gca} + \sigma^2_{tgca})}.$$

This measures the degree of GXE at the additive level ($r_B = 1$ implies no GXE, while $r_B = 0$ implies that there is no consistent rust resistance performance of families across test sites).

From the results of the paired-test analyses it was also possible to estimate additive genetic correlations, in this case age-age genetic correlations, from each test-pair where both tests were measured at age_1 and age_2:

$$r_{g_{(age1,\,age2)}} = \frac{\sigma_{gca_{(age1,\,age2)}}}{\sqrt{(\sigma^2_{gca_{(age1)}}) \cdot (\sigma^2_{gca_{(age2)}})}}$$

where the additive genetic covariance between two ages ($\sigma_{gca(age1,\ age2)}$) is the quadratic average covariance for that test-pair, i.e.,

$$\sqrt{\sigma_{gca_{(age1,\,age2)}} \cdot \sigma_{gca_{(age2,\,age1)}}}$$

and $\sigma^2_{gca(age1)}$ and $\sigma^2_{gca(age2)}$ are estimated from the same test pair where data are the same age in both tests. This requires both tests to have been measured at both age_1 and at age_2.

Linear regression analysis was used in an attempt to develop predictive equations for each of the genetic parameters described above, using mean rust incidence of each test, and various transformations and combinations of mean rust incidence as regressors. All final models were required to a) be significant at $p=0.0001$ (F-test) with no consequential lack of fit, b) have all independent variables significant at $p=0.01$ (t-test), and c) to be biologically viable across the entire range of possible regressors. If such models could not be identified then simple unweighted average parameters across all tests or test pairs are reported.

Results and Discussion

Useful regression models (Table 4-1) were identified for biased and unbiased heritability of rust resistance (h_b^2 and h^2, from single- and paired-sites with values of the coefficient of

determination, R^2, of 0.33 and 0.43 respectively), and type B genetic correlation between two sites (r_B with R^2=0.18). However no models adequately predicted the proportion of dominance variance (either biased or unbiased estimates) or the age-age genetic correlation (r_g). Where possible, an approximate test for lack of fit was constructed by subdividing the independent variable into classes (Weisberg 1985, p.95), and plots were used to assess the validity of the models. For all models reported no important lack of fit was detected even though the R-square values reported are generally only modest.

Table 4-1. Regression models to predict single-site biased heritability (h_b^2), paired-site unbiased heritability (h^2), and type B genetic correlations (r_B), for resistance to fusiform rust developed from 171 slash pine full-sib tests. (Regressors in models: R = mean rust incidence at one site; RSUM = sum of the mean rust incidence at two sites; RAVG = RSUM/2; RDIFF = absolute difference in mean rust incidence at two sites.)

Dependent Variable	Age	Regression Model	R^2	n
Heritability of Rust in one test (h_b^2)	All (4-15)	$\hat{h}_b^2 = 0.000156 + 0.001295(R)^{1.5} - 0.000115(R)^{2.0}$	0.33	362
Heritability of Rust in a pair of tests (h^2)	All (4-15)	$\hat{h} = -0.022301 + 0.003228(RAVG)$	0.43	955
Type B genetic correlation of rust between two tests (r_B)	All (4-15)	$\hat{r}_B = 0.405322 + 0.008709(RSUM) - 0.00003162(RSUM)^2 - 0.007124(RDIFF)$	0.18	908

Age, mating design (diallel or factorial), and the type of cross (crosses between initial first-generation parents or crosses with rust-free parents) were not found to be significant (p-value ≤ 0.05) in any model attempted. Therefore all data regardless of age, mating design and cross-type were used together for both single- and paired-site models. Parameters were successfully modelled using simple functions of the mean level of rust incidence in a test (R), the sum of the mean rust incidence in two tests (RSUM = $R_1 + R_2$), the average rust incidence in two tests (RAVG = RSUM/2), or the absolute difference in the mean rust incidence of the two tests (RDIFF = $|R_1 - R_2|$), expressed in percentage units. The relatively low R-square of the model

predicting the type B genetic correlation (Table 4-1) was reasonable given the difficulty of estimating genetic correlations (Falconer 1989, p.317) and experimental error associated with the estimation of parental performance. All the predictive models presented are believed to have greater utility than either the parameter estimates themselves, or simple average parameters which take little or no account of difference in rust infection.

Heritability

Biased heritability (h_b^2) was predicted by a quadratic function of the mean rust in the test (Table 4-1). The model presented was better than simply using the mean rust incidence and its square: this alternative model had significant lack of fit, and a lower R-square value. The predicted maximum biased heritability is 0.195 when the mean rust incidence equals 71.5% (Figure 4-1). The model predicting unbiased heritability (h^2) from paired tests was a simple linear function of the average rust incidence in the two tests (Table 4-1), and unlike the predicted biased heritability no significant quadratic effect was detected (Figure 4-1). All models for h^2 investigated which included a quadratic term, showed little deviation from a straight line, and did not improve the R-square. It is possible that the lack of a quadratic trend in the h^2 model was related to the very limited amount of data which was available for the upper range of average rust infection levels. The maximum value of RAVG observed was 87.5%, and there were only 29 out of a total of 955 test-pairs where RAVG exceeded 75%. Since pairs of connected tests with high levels of rust in both tests were relatively rare, Figure 4-1 is only plotted for average rust incidence between 0 and 80%. If additional data had been available from test-pairs with high levels of rust infection, it is possible that the predicted h^2 may have declined in a manner similar to the predicted h_b^2 above 70% rust infection levels. However in defense of the linear model presented it should be noted that the predicted h^2 at RAVG = 100% is 0.30, which is similar to

realized heritability estimates, $h_r^2 \cong 0.35$ at rust infection levels exceeding 90% (Hodge *et al.* 1990). Within the range of 0-60% rust the predicted h^2 was consistently lower than the predicted h_b^2 (Figure 4-1), which would be expected if there were any significant additive genetic-by-environment interaction. A check on the validity of these models can be obtained in the following manner: i) by definition $h_b^2 = h^2 \div r_B$, ii) at an average rust infection level of 50%, h_b^2 is predicted to be 0.171, h^2 is predicted to be 0.139, and r_B is predicted to be 0.823 (using RDIFF = 19.2% — the mean RDIFF over all 955 observations), and, thus, iii) $\hat{h}_b^2 = 0.139 \div 0.823 = 0.169$. This second estimate of h_b^2, 0.169 is very close to that predicted from the model in Table 4-1, i.e., $h_b^2 = 0.171$. Similarly, if we predict h_b^2 in this same manner in 2% increments in mean rust incidence from 20-100% (below 20% an RDIFF of 19.2% makes little sense), within this range the average difference between the two predictions was 0.018 and is never greater than 0.025. But above 66% rust infection levels, the absolute difference increases rapidly to over 0.16 at 100% rust infection. Thus it would appear that the model for unbiased heritability is most reliable only for RAVG values below 66%.

Proportion of Dominance

Attempts to develop predictive models for biased and unbiased estimates of the proportion of dominance (d_b^2 and d^2 respectively), and for the ratio of additive to dominance variance (h^2/d^2), were fruitless. No model with an R-square exceeding 0.09 could be found for d_b^2, d^2, or h^2/d^2. Analysis of variance revealed that neither age classes, rust classes, nor the interaction between age and rust classes was significant in the case of d_b^2. For paired-site analyses age classes were also not significantly different and test pairs can not be readily classified on the basis of rust incidence. In the absence of reliable models, or meaningful classification variables, the unweighted mean proportion of dominance was estimated to be 0.087 and 0.053 for d_b^2 and d^2,

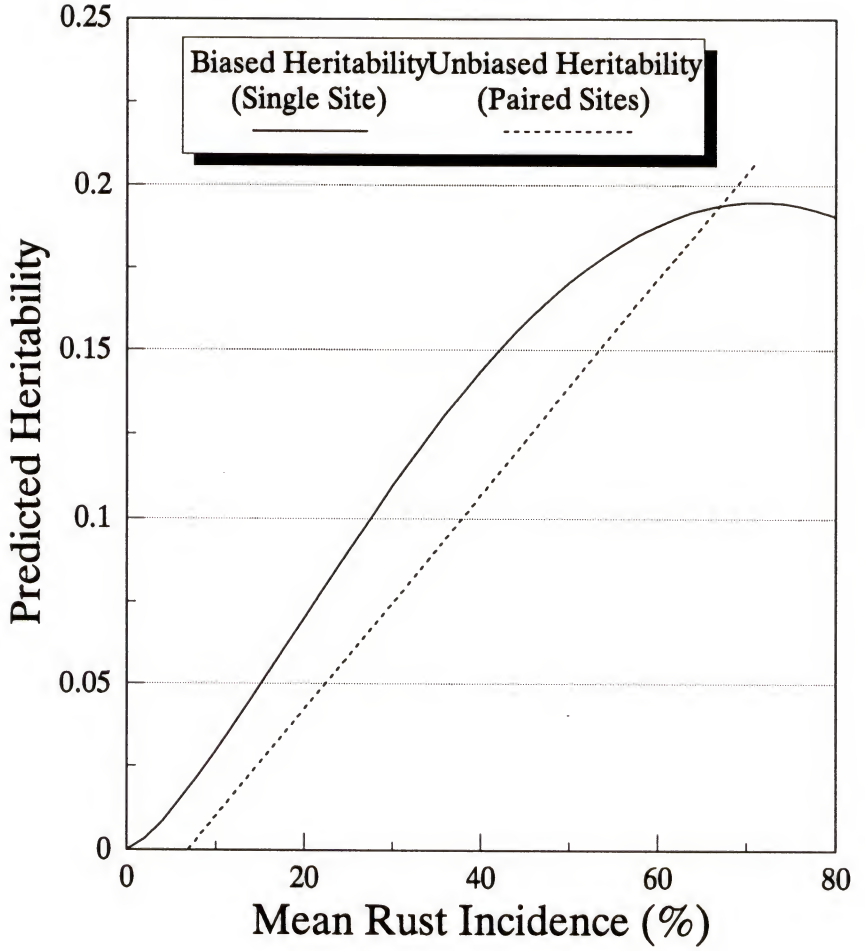

Figure 4-1. Predictive models for the single-site biased heritability (h_b^2) and paired-site unbiased heritability (h^2) of resistance to fusiform rust in slash pine. Models developed using data from 171 full-sib slash pine progeny tests measured between 4 and 15 years after field planting. ($\hat{h}_b^2$ = 0.00156 + 0.001295(R)$^{1.5}$ + 0.000115(R)^2, and $\hat{h}$ = -0.022301 + 0.003228(RAVG) with R-squares of 0.33 and 0.43, respectively. R = mean rust infection at one site, and RAVG = average rust infection at two sites.)

respectively, in tests where the mean rust incidence was between 10 and 90 percent (Table 4-2). Tests with extreme levels of rust were excluded because variance is low in such tests (Sohn and Goddard 1979). Average biased estimates were higher than unbiased estimates indicating the presence of dominance-by-environment interaction which is about 64% of the dominance variance.

If the proportion of dominance in fusiform rust resistance of slash pine is constant across levels of rust incidence and age classes between 4 and 15 years, as these data indicate, then the ratio of additive to dominance variance ($\sigma_A^2/\sigma_D^2 = h^2/d^2$) can be estimated by dividing the predicted heritability by the mean proportion of dominance. For single tests within the limits of 10-90% rust incidence, σ_A^2/σ_D^2 has a maximum of 2.2 at a rust incidence of 71.5%, and a minimum of 0.34 at 10% rust.

Table 4-2. Mean proportion of dominance in the fusiform rust resistance of slash pine, estimated from single (biased, d_b^2) and paired (unbiased, d^2) full-sib tests. (Means include only tests where the mean rust incidence is between 10% and 90%)

Analysis Type	Number of Tests or Test Pairs	Proportion of Dominance (± standard error of mean)
Single-tests	307	0.087 ± 0.007
Paired-tests	674	0.053 ± 0.003

Genetic Correlations

Type B genetic correlation for fusiform rust resistance between two tests was modelled as a function of the sum of the mean rust incidence in the two tests (RSUM) and absolute difference in the mean rust levels in the two tests (RDIFF), with an R-square of 0.18 (Table 4-1). Although the R-square for this model was relatively small, the model is biologically feasible. The predicted r_B was at a maximum when the two tests had the same mean rust infection level, and

decreased as the difference in the rust incidence between the two tests increased (Figure 4-2). When rust levels in both tests were intermediate (30-70%), type B genetic correlations were predicted to exceed 0.67. Shelbourne (1972) suggested that when the variance of GXE is at least half the size of the additive variance (i.e., Type B = 0.67), breeding strategies should be modified to address this interaction, or potential genetic gains will be compromised. Therefore, at intermediate levels of rust infection, GXE interaction would appear to be of little importance for rust resistance in slash pine.

No significant regression models were found for genetic correlations (r_B) of rust resistance at two different ages between 5 and 14 years. To prevent estimates well outside the parameter space from unduly influencing average parameter estimates, individual estimates of r_B exceeding 1.3 were set to 1.3. Unweighted average age-age genetic correlations estimated from tests where the mean rust incidence in both tests was between 10 and 90% are not significantly different from 1.0 (Table 4-3). These very strong age-age correlations are reasonable in view of the fact that for rust resistance no other genetic parameter showed any relationship to age.

Table 4-3. Average age-age genetic correlation (r_g) of resistance to fusiform rust in slash pine, estimated from full-sib tests ($\pm$ the standard error of the means). Individual estimates were bounded at a maximum of 1.3, and the means only include data from test-pairs where the mean rust infection in both tests was between 10% and 90%.

Age 1 (Years)	Age 2 (Years)	Number of Test Pairs	Genetic Correlation ($r_g \pm$ s.e.)
Five	Eight	144	0.911 $\pm$ 0.034
	Eleven	54	0.951 $\pm$ 0.031
	Fourteen	8	1.001 $\pm$ 0.016
Eight	Eleven	45	0.960 $\pm$ 0.083
	Fourteen	9	0.800 $\pm$ 0.120
Eleven	Fourteen	2	0.980 $\pm$ 0.022

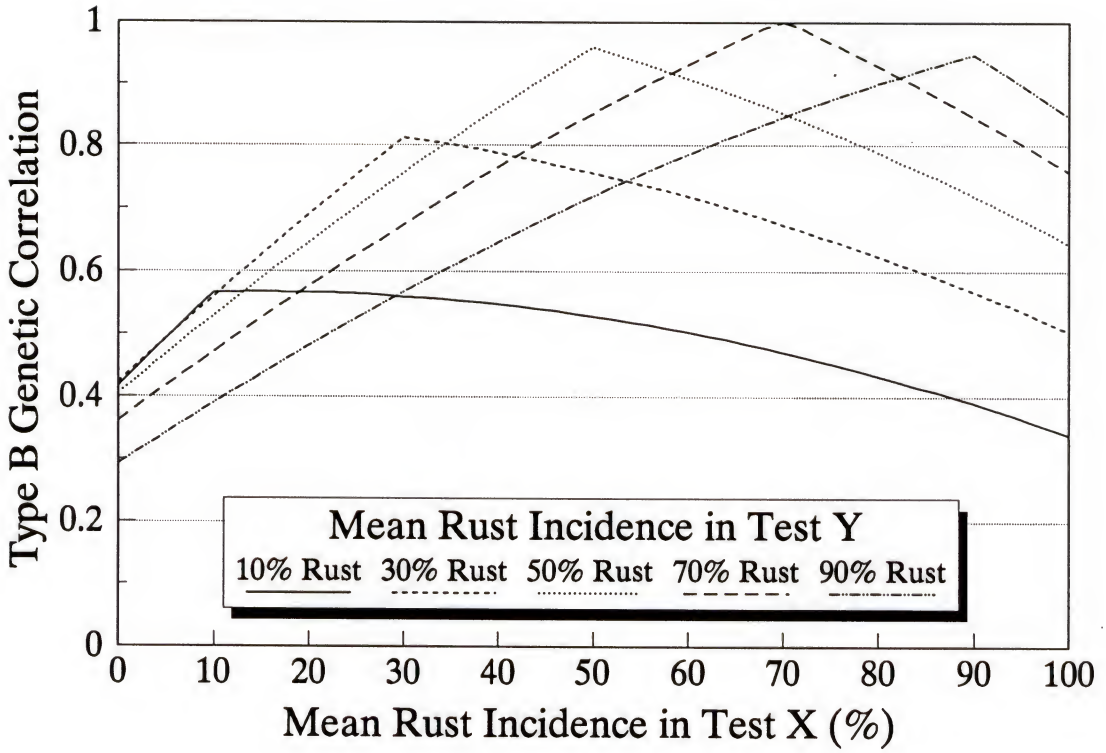

Figure 4-2. Predicted type B genetic correlation (r_B) for resistance to fusiform rust, between two slash pine progeny tests (Tests X and Y). Predictive model developed using data from 171 full-sib slash pine progeny tests between 4 and 15 years after field planting. ($\hat{r}_B$ = 0.405322 + 0.008709(RSUM) + 0.00003162(RSUM)^2 - 0.007124(RDIFF); R-square = 0.18 and p-value for model = 0.0001; RSUM = sum of the mean rust incidence at two sites, RAVG =RSUM/2, and RDIFF = absolute difference in the mean rust incidence at two sites)

Implications for Breeding Strategy

For most quantitative traits breeders attempt to identify a minimum selection age which will maximize economic and genetic gains from improvement work. Data presented here show that, for resistance to fusiform rust in slash pine, the mean infection level is a more important determinant of future gains than is age (age was not important in determining any genetic parameter estimated, nor were age-age genetic correlations significantly different from one). When selecting for resistance to fusiform rust infection, provided the trees have been exposed to at least moderate infection levels, there appears to be no benefit in delaying selection beyond five years. Also, because infection levels are a primary determinant of heritability, and hence genetic gain, it is important that tests be established in areas where they are likely to be exposed to the pathogen, i.e., moderate to high rust hazard sites.

Genotype-by-environment interaction for fusiform rust resistance was detected; however, it does not appear to be of any major consequence. At intermediate levels of rust infection levels (30-70%) type B genetic correlations are always sufficiently high to negate concerns about GXE reducing potential gains. Also, as long as selection for rust resistant individuals is restricted to sites with high rust infection levels, then heritability is high (Figure 4-1) and type B genetic correlations with other test are high (Figure 4-2).

The proportion of dominance in slash pine was not found to change with age or rust infection levels, but to be constant at relatively low levels across all parameters examined. Dominance variance is, however, only predicted to be important where the mean infection level is low, i.e., where the amount additive variance is low. However because resistance to rust is of little importance on sites with low levels of fusiform rust, dominance variance is not regarded to be of any significant importance to the breeding of resistant genotypes for future deployment.

These results reaffirm assumptions about the relative unimportance of dominance variance underlying the current CFGRP breeding strategy (White *et al* 1993) and give no strong reasons for major changes to this strategy which is based on recurrent selection for general combining ability. However the results do reinforce the need to establish progeny tests on sites where rust infection is likely to be moderate to high, if substantial gains in rust resistance are to be achieved in future generations, and that selection 'age' should be determined by mean rust infection levels rather than the age of the test.

Conclusions

Significant regression models for fusiform rust resistance were identified to predict biased and unbiased heritability estimates, and the type B genetic correlation between two sites as functions of the mean rust incidence in the tests. The maximum predicted single-site (biased) heritability is 0.195 at 71.5% rust infection. A simple linear relationship was modelled between the unbiased (paired-site) heritability of rust resistance and the mean rust infection level. However data where both tests encountered high levels of rust infection were rare, and the model for unbiased heritability appears to be most reliable for mean rust infection levels below 66%. Type B genetic correlations indicated the existence of GXE; however, at intermediate to high levels of rust infection (>50%), GXE does not appear to be of any real importance in the rust resistance of slash pine. The proportion of dominance was not found to be related to rust incidence or age, and average estimates of 0.087 and 0.053 were obtained for d_b^2 and d^2, respectively. Dominance variance is predicted to be relatively small compared to the additive variance except at low levels of rust infection. Therefore, dominance variance is not considered to be of any significant importance to the improvement of rust resistance in slash pine. Age was not a significant factor affecting the estimation of any genetic parameter examined, and all age-age genetic correlations

for rust resistance were close to 1.0. Therefore age appears to unimportant in determining the quantitative genetics of rust resistance in slash pine. Breeding strategies that concentrate on the utilization of additive variance, and which restrict testing and selection efforts to sites with at least 50% rust incidence, are expected to be the most successful strategies for the improvement of fusiform rust resistance in slash pine. These conclusions reinforce the current CFGRP breeding strategy for slash pine.

CHAPTER 5
CONCLUSIONS

In any successful, advanced-generation tree improvement program detailed knowledge of the quantitative genetics is very important. By utilizing data from 171 full-sib tests of slash pine established in the southeastern United States by the Cooperative Forest Genetics Research Program, it has been possible to i) investigate the relationship between approximate and empirical estimates of the variance of the heritability estimates of tree volume and resistance to fusiform rust infection, ii) obtain reliable average genetic parameter estimates for tree volume, and to investigate the relationship between these parameters and test characteristics such as age and site index, and iii) to model the relationship between mean incidence levels of fusiform rust and some genetic parameters of rust resistance in slash pine. This study is distinguished from previous work by the magnitude of the data which was used, the use of Restricted Maximum Likelihood estimation procedures, and the use of full-sib data which allows the estimation of the importance of dominance variance. Previous work in most forest tree species has concentrated on a small number of tests, often used sub-optimal estimation procedures, and frequently used half-sib material from which no estimate of dominance variance can be obtained.

Two approximate methods, Taylor series and Dickerson (1969), for estimating the variance of heritability estimates (both biased estimates from single site analyses and unbiased estimates obtained from paired site analyses, for both volume and rust resistance), proved to yield estimates that were of a similar magnitude to each other and to an empirical estimate of the variance of heritability estimates. However for biased heritability both approximate methods

tended to underestimate the empirical estimate of the variance. Irrespective of the trait, test size, or whether REML of ANOVA-based estimation of the variance components was used, the Dickerson method was slightly more conservative than the Taylor series approximation, and hence are generally closer to the empirical estimates. There was some evidence of small-sample bias when using the asymptotic variances and covariances of REML estimates in the formulae for the Taylor and Dickerson approximations.

Overall, the Taylor series and Dickerson approximations provided essentially the same information about the variance of heritability estimates, and both were reasonably good approximations to empirical estimates of the variance. Of these two methods, the Dickerson approximation seems to be slightly superior because it is more conservative and considerably simpler to calculate. Even if researchers opt to use the Taylor series approximation, the Dickerson method is a quick method to verify the accuracy of their calculations.

Genetic parameters (heritability, the proportion of dominance, type B genetic correlations, and age-age genetic correlations) were estimated for both tree volume and resistance to fusiform rust at multiple ages between 4 and 15 years in slash pine. Attempts to model these genetic parameter estimates using age and test characteristics as regressors proved to be unsuccessful for tree volume, but for the heritability and type B genetic correlations of rust resistance it was possible to develop significant and useful predictive models. Average parameter estimates were obtained for all genetic parameters for which models could not be identified.

The heritability of volume in slash pine was found to be relatively low compared to published estimates in other tree species. Narrow sense heritability of volume at five years was 0.07, and increased to around 0.12 at 11 and 14 years of age. These estimates are similar to previous estimates from the same breeding population but using open-pollinated test data (Hodge and White 1992). Average estimates of the proportion of dominance in tree volume (of this slash

pine population) were fairly high. At 5 years of age there was about 1.6 times as much additive variance as dominance variance which increased to over two times at 11 or more years of age. Further work is required to assess the economic merit of pursuing dominance variance to obtain additional genetic gain, but given the apparent declining importance of dominance in volume with increasing age, dominance may prove to be relatively unimportant at rotation age. Estimates of type B genetic correlations for volume in slash pine increased from 0.6 at 5 years to 0.8 at 14 years, indicating that genotype-by-environment interaction diminished as the tests aged; these estimates are of a similar magnitude to previous estimates from open-pollinated tests (Hodge and White 1992). Estimates of age-age genetic correlations for volume growth were modelled successfully as a function of the natural logarithm of the age ratio (Lambeth 1980), and were slightly lower than estimates from open-pollinated tests (Hodge and White 1992). They indicated that age 11 data was essentially the same as that from 14 years, but that selection at younger ages would result in a loss of genetic gain.

For resistance to fusiform rust in slash pine, significant regression models were developed to predict biased and unbiased heritability, and type B genetic correlation between two sites, as functions of the mean rust incidence in the tests. However for the other genetic parameters estimated (the proportion of dominance, and age-age genetic correlation) no useful models could be identified. Biased heritability of rust resistance was predicted to maximize at $h_b^2 = 0.195$, when the mean rust infection was 71.5%. The model for unbiased heritability estimates (from paired tests), was a linear function of the mean rust infection level, but was most reliable only for rust infection levels between 0 and 66% because there were relatively few observations where both tests encountered high rust infection levels. Within the range 0-66% rust, the predicted biased heritability of rust resistance exceeded the predicted unbiased heritability by an amount predictable from the type B genetic correlation. The predictive model for type B genetic

correlations indicated the existence of genotype-by-environment interaction in rust resistance of slash pine; however, at intermediate to high levels of rust infection genotype-by-environment interaction was of little importance. The proportion of dominance was not found to be related to rust incidence, test age, or any other test parameter, and average estimates of 0.088 and 0.053 were obtained for biased (d_b^2) and unbiased (d^2) estimates, respectively. The importance of dominance variance relative to additive variance was predicted to be small, except at low levels of rust infection. However, as resistance to fusiform rust is of little importance when infection levels are low, dominance variance is apparently of no importance to the continued improvement of rust resistance in slash pine. Test age was not a significant factor affecting the estimation of any of the genetic parameters examined, and all age-age genetic correlations were close to 1.0. Hence age seems to be unimportant in determining the quantitative genetics of rust resistance in slash pine.

Although current breeding strategies for slash pine (White *et al.* 1993) were developed in the absence of detailed estimates of the importance of dominance variance in tree volume and rust resistance, the results of this work generally confirm the breeding strategy currently in place. For tree volume, there may be some advantages to pursuing additional genetic gain from dominance variance in the elite, most superior, portion of the breeding population; however, the potential gains must be weighed carefully against the additional costs. For rust resistance in slash pine, the current strategy which concentrates on gains from additive variance, in tests with moderate to high levels of rust infection, would appear to be near optimal.

APPENDIX

SOME PUBLISHED HERITABILITY ESTIMATES FOR *PINUS* SPECIES

1. Slash Pine (*Pinus elliottii* var *elliottii*)

Source	Age (Yrs)	Volume		Diameter		Height	
		h_b^2	h^2	h_b^2	h^2	h_b^2	h^2
Dorman and Squillace (1974)	—	0.16 to 0.35		-0.22 to 0.58		0.03 to 0.37	
Sluder (1986)	21	0.15		0.16		0.00	
	22	0.16		0.19		0.20	
Cotterill *et al.* (1987)	8	0.37		0.36		0.37	
Falkenhagen (1989)	5	0.34		0.35		0.39	
	8	0.41		0.40		0.38	
Hodge and White (1992)	5	0.14	0.08	0.13	0.06	0.14	0.08
	10	0.23	0.16	0.23	0.15	0.20	0.13
	15	0.26	0.16	0.24	0.16	0.23	0.12
Pswarayi (1993)	5	0.51	0.48	0.43	0.44	0.44	0.39
	10	0.49	0.40	0.43	0.39	0.40	0.27
	15	0.55	0.29	0.54	0.31	0.37	0.10

2. Loblolly Pine (*Pinus taeda*)

Source	Age (Yrs)	Volume		Diameter		Height	
		h_b^2	h^2	h_b^2	h^2	h_b^2	h^2
Matziris and Zobel (1973)	5	0.28		0.20		0.44	
Bridgwater *et al.* (1983)	10		0.18		0.10		0.28
Lambeth *et al.* (1983)	5	0.38				0.45	
	13	0.31				0.41	
	16	0.43				0.66	
	20	0.47				0.61	
Foster (1986)	3	0.09		0.15		0.45	
	5	0.09		0.05		0.17	
	8	0.00		0.00		0.13	
	10	0.03		0.03		0.28	
	15	0.06		0.04		0.41	
Talbert (1984)	4						0.05
	6		0.03		0.05		0.07
	8		0.08		0.41		0.17
	10		0.11		0.12		0.18
	13		0.12		0.13		0.19
	15	0.15	0.11	0.15	0.14	0.28	0.20
Balocchi (1990) and Balocchi *et al.* (1993)	3						0.04
	5						0.04
	4						0.07
	10						0.12
	15						0.23
	26						0.18

3. Other *Pinus* Species

Species	Source	Age (Yrs)	Volume		Diameter		Height	
			h_b^2	h^2	h_b^2	h^2	h_b^2	h^2
P. radiata	Johnson & Burdon (1990)	4	0.20		0.19		0.23	
	Cotterill *et al.* (1987)	8	0.49		0.09		0.12	
	Cotterill & Dean (1988)	6 10					0.36 0.32	
P. patula	Barnes *et al.* (1992)	5 5 5	0.27 0.36	0.18 0.17			0.27 0.19	0.13 0.08
	Ladrach and Lambeth (1991)	7	0.49	0.38			0.44	0.38
P. caribaea var *hondurensis*	Dean *et al.* (1986)	4 7 10			0.41 0.48 0.52		0.27 0.29 0.23	
	Woolaston *et al.* (1990)	5			0.30		0.20	
P. palustris	Snyder & Namkoong (1978)	7	0.03				0.19	
	Sluder (1986)	21	0.23		0.27		0.30	
P. echinata	Tauer & McNew (1985)	4 10		0.72		0.59		0.34 0.61
P. banksiana	Riemenschneider (1988)	3 5 7					0.20 0.21 0.20	

REFERENCE LIST

Adams, W.T. and Joyce, D.G. 1990. Comparison of selection methods for improving volume growth in young coastal Douglas-fir. Silvae Genet. 39(5-6):219-226

Adams, W.T., Neale, D.B. and Loopstra, C.A. 1988. Verifying controlled crosses in conifer tree-improvement programs. Silvae Genet. 37(3-4):147-152

Ades, P.K., Simpson, J.A., Eldridge, K.G. and Eldridge, R.H. 1992. Genetic variation in susceptibility to Dothistroma needle blight among provenance and families in Pinus muricata. Can. J. For. Res. 22:1111-1117

Anderson, R.L., McClure, J.P., Cost, N. and Uhler, R.J. 1986. Estimating fusiform rust losses in five southeast states. South. J. Appl. For. 10:237-240

Bailey, R.L. and Brooks, J.R. 1994. Determining site index and estimating timber volumes without measuring heights. Southern J. Appl. For. 18:15-18

Baker, R.J. 1986. Selection Indices in Plant Breeding. CRC Press, Inc., Boca Raton, FL, 218p.

Balocchi, C.E. 1990. Age trends of genetic parameters and selection efficiency for loblolly pine (Pinus taeda L.). Ph.D. Thesis, Dept. For., North Carolina State Univ., Raleigh, NC, 94p.

Balocchi, C.E., Bridgwater, F.E., Zobel, B.J. and Jahromi, S. 1993. Age trends in genetic parameters for tree height in nonselected populations of loblolly pine. For. Sci. 39:231-251

Banks, B.D., Mao, I.L. and Walter, J.P. 1985. Robustness of the restricted maximum likelihood estimator derived under normality as applied to data with skewed distributions. J. Dairy Sci. 68:1785-1792

Barnes, R.D., Mullin, L.J. and Battle, G. 1992. Genetic control of fifth year traits in Pinus patula Schiede and Deppe. Silvae Genet. 41:242-248

Becker, W.A. 1975. Manual of Quantitative Genetics. Washington State Univ. Press, Pullman, WA, 170p.

Borralho, N.M.G., Cotterill, P.P. and Kanowski, P.K. 1992. Genetic parameters and gains expected from selection for dry weight in Eucalyptus globulus ssp. globulus in Portugal. For. Sci. 38:80-94

Bridgwater, F.E., Talbert, J.T. and Jahromi, S. 1983. Index selection for increased dry weight in young loblolly pine population. Silvae Genet. 32:157-161

Byram, T.D. and Lowe, W.J. 1986. General and specific combining ability estimates for growth in loblolly pine. In: Proc. IUFRO Symp. on Breeding Theory, Progeny Testing and Seed Orchards. Oct 13-17, Williamsburg, VA, pp. 352-360

Bulmer, M.G. 1980. The Mathematical Theory of Quantitative Genetics. Clarendon Press, Oxford, UK 255p.

Burdon, R.D. 1977. Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genet. 26:168-175

Carson, S.D. 1989. Selecting Pinus radiata for resistance to dothistroma needle blight. NZ J. For. Sci. 19:3-21

Cockerham, C.C. 1963. Estimation of genetic variances. In: Statistical Genetics and Plant Breeding. R.E. Hanson and H.F. Robinson (Eds.). NAS-NRC Pub. 982, Washington DC, pp. 53-93.

Comstock, R.E. and Moll, R.H. 1963. Genotype-environment interactions. In: Statistical Genetics and Plant Breeding. R.E. Hanson and H.F. Robinson (Eds.). NAS-NRC Pub. 982, Washington DC, pp. 169-194.

Cornelius, J. 1994. Heritabilities and additive genetic coefficients of variation in forest trees. Can. J. For. Res. 24:372-379

Cotterill, P.P. 1987. Short note: on estimating heritability according to practical applications. Silvae Genet. 36:46-48

Cotterill, P.P. and Dean, C.A. 1988. Changes in the genetic control of growth in radiata pine to 16 years and efficiencies of early selection. Silvae Genet. 37:138-146

Cotterill, P.P., Dean, C.A. and van Wyk, G. 1987. Additive and dominance genetic effects in Pinus pinaster, P.radiata and P.elliottii and some implications for breeding strategy. Silvae Genet. 36:221-232

Dean, C.A., Cotterill, P.P. and Eisemann, R.L. 1986. Genetic parameters and gain expected from selection in Pinus caribaea var. hondurensis in northern Queensland, Australia. Silvae Genet. 35:229-236

De Souza, S.M. 1990. Greenhouse testing of slash pine for resistance to fusiform rust. Ph.D. Dissertation, For. Res. Conserv., Univ. Florida, Gainesville, FL, 92p.

De Souza, S.M.; White, T.L.; Hodge, G.R.; Schmidt, R.A. 1991. Genetic parameter estimates of greenhouse traits of slash pine artificially inoculated with fusiform rust fungus. For. Sci. 37:836-848

Dickerson, G.E. 1962. Implications of genetic-environmental interaction in animal breeding. Animal Production 4:47-64

Dickerson, G.E. 1969. Techniques for research in quantitative animal genetics In: Techniques and Procedures in Animal Science Research. Amer. Soc. Anim. Sci., Albany, NY, pp. 36-79.

Dieters, M.J.J., Woolaston, R.R. and Nikles, D.G. 1990. Internode length of hoop pine: genetic parameters and prospects for developing a long-internode breed. NZ J. For. Sci. 20:138-147

Dingus, R.J. and Griggs, M.M. 1975. Rust-free slash pines surviving artificial inoculation potentially useful for resistance breeding. For. Sci. 21:275-277

Dorman, K.W. and Squillace, A.E. 1974. Genetics of slash pine. USDA For. Serv. Res. Pap. WO-20, 20p.

Falconer, D.S. 1989. Introduction to Quantitative Genetics. Longman Scientific and Technical, Longman Group, Hong Kong, 438p.

Falkenhagen, E.R. 1989. Relationships between some genetic parameters and test environments in open-pollinated families of Pinus elliottii in South Africa. Theor. Appl. Genet. 77: 857-866

Foster, G.S. 1986. Trends in genetic parameter estimates with stand development and their influence on early selection of volume growth in loblolly pine. For. Sci. 32:944-959

Giesbrecht, F.G. 1983. An efficient procedure for computing MINQUE of variance components and generalized least squares estimates of fixed effects. Commun. Statist. Theor. Meth. 12:2169-2177

Goddard, R.E. 1980. The University of Florida cooperative forest genetics research program. In: Proc. 15th North Amer. Quant. For. Genet. Group Workshop, Coeur d'Alene, ID., Aug. 6-8, pp. 31-42.

Goddard, R.E. and Schmidt, R.A. 1979. Relative geographic stability of resistance to fusiform rust of selected slash pine families. In: Proc. South. For. Tree Imp. Conf., June 19-21, Mississippi State Univ., pp. 99-107

Goddard, R.E., Schmidt, R.A and Vande Linde, F. 1975. Effect of differential selection pressure on fusiform rust resistance in phenotypic selections of slash pine. Phytopath. 65:336-338

Goddard, R.E. and Strickland, R.K. 1968. Volume and weight tables for five-year-old plantation grown slash pine. School of Forest Resources, Rep. No. 14, University of Florida, Gainesville, FL, 8p.

Goodman, A.W. 1969. Analytical Geometry and the Calculus. Collier-Macmillan Canada, Ltd., Toronto, Ontario (Second Edition), 819p.

Griffing, B. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. Aust. J. Biol. Sci. 9:463-493

Hallauer, A.R. and Miranda, J.B. 1988. Quantitative Genetics in Maize Breeding. Iowa State Univ. Press, Ames IA, (Second Edition), 468p.

Hartley, H.O. and Rao, J.N.K. 1967. Maximum-likelihood estimation for the mixed analysis of variance model. Biometrika 54:93-108

Harville, D.A. 1977. Maximum likelihood approaches to variance component estimation and to related problems. J. Amer. Stat. Assoc. 72:320-338

Hendrickson, J.A. 1976. Selection program for improved rust resistance of slash pine. In: Eighteenth Progress Report Cooperative Forest Genetics Research Program. R.E. Goddard, D.L. Rockwood, H.R. Kok, C.A. Hollis and J.A. Hendrickson (Eds.). Dept. For., Univ. Florida, Gainesville, FL, 20p.

Hodge, G.R., Schmidt, R.A. and White, T.L. 1990. Substantial realized gains from mass selection of fusiform rust-free trees in highly infected stands of slash pine. Southern J. Appl. For. 14:143-146

Hodge, G.R. and White, T.L. 1986. Relationship of fusiform rust and growth traits in open-pollinated slash pine progeny tests. In: Proc. IUFRO Symp. on Breeding Theory, Progeny Testing and Seed Orchards. Oct 13-17, Williamsburg, VA, pp. 469-481

Hodge, G.R. and White, T.L. 1992. Genetic parameter estimates for growth traits at different ages in slash pine and some implications for breeding strategy. Silvae Genet. 41:252-262

Hodge, G., White, T. and Powell, G. 1991. Thirty-Third Progress Report, Cooperative Forest Genetics Research Program, Dept. For., Univ. Florida, Gainesville, FL. 55p.

Hodge, G.R., White, T.L., Powell, G.L. and de Souza, S.M. 1989. Predicting genetic gains from one generation of slash pine tree improvement. South. J. Appl. For. 13:51-56

Hodge, G., White, T., Powell, G., Rockwood, D. and Parker, S. 1994. Thirty-sixth Progress Report, Cooperative Forest Genetic Research Program, Dept. For., Univ. Florida, Gainesville, FL., 29p.

Huber, D.A. 1993. Optimal mating designs and optimal techniques for analysis of quantitative traits in forest genetics. Ph.D. Dissertation, For. Res. Conserv., Univ. Florida, Gainesville, FL, 151p.

Johnson, G.R. and Burdon, R.D. 1990. Family-site interaction in Pinus radiata: implications for progeny testing strategy and regionalised breeding in New Zealand. Silvae Genet. 39:55-62

Kaplan, W. 1952. Advanced Calculus. Addison-Wesley Publishing Company, Inc., Reading, MA.

Kempthorne, O. 1957. An Introduction to Genetic Statistics. John Wiley and Sons, Inc., New York, NY, 545p.

Kendall, M.G., Stuart, A. and Ord, J.K. 1987. Kendall's Advanced Theory of Statistics. Volume 3, Design and Analysis, and Time Series. Oxford Univ. Press, New York, 780p.

Khattree, R. and Gill, D.S. 1988. Comparison of some estimates of variance components using Pitman nearness criterion. In: Amer. Stat. Assoc., Proc. Stat. Computing Sect., New Orleans, LA, pp. 133-136.

Khuri, A.I. and Sahai, H. 1985. Variance components analysis: a selective literature survey. International Statistical Review 53:279-300

King, J.N., Yeh, F.C., Heaman, J.C.H. and Dancik, B.P. 1988. Selection of wood density and diameter in controlled crosses of coastal Douglas-fir. Silvae Genet. 37:152-157

Kiss, G. and Yeh, F.C. 1988. Heritability estimates for height for young interior spruce in British Columbia. Can. J. For. Res. 18:158-162

Knapp, S.J., Bridges, W.C. and Yang, M. 1989. Non-parametric confidence interval estimators for heritability and expected selection response. Genetics 121:891-898

Knapp, S.J., Stroup, W.W. and Ross, W.M. 1985. Exact confidence intervals for heritability on a progeny mean basis. Crop Sci. 25:192-194

Kraus, J.F. 1973. Estimates of general and specific combining ability for height and rust resistance from single crosses of slash pine. Silvae Genet. 22:121-124

Ladrach, W.E. and Lambeth, C. 1991. Growth and heritability estimates for a seven-year-old open-pollinated Pinus patula progeny test in Columbia. Silvae Genet. 40:169-173

Lambeth, C.C. 1980. Juvenile-mature correlation in Pinaceae and implication for early selection. For. Sci. 26:571-580

Lambeth, C.C., van Buijtenen, J.P., Duke, S.D. and McCullough, R.B. 1983. Early selection is effective in 20-year-old genetic tests of loblolly pine. Silvae Genet. 32:210-215

Lowe, W.J. and Greene, T.A. 1990. Geographic variation in specific gravity and fiber length of green ash (Fraxinus pennsylvanica Marsh.) in east Texas. Silvae Genet. 39:194-198

Matziris, D.I. and Zobel, B.J. 1973. Inheritance and correlations of juvenile characteristics in loblolly pine (Pinus taeda L.). Silvae Genet. 22:38-45

McKeand, S.E., Foster, G.S. and Bridgwater, F.E. 1986. Breeding systems for pedigree-controlled production populations of loblolly pine. In: Advanced Generation Breeding of Forest Trees, Lous. Agric. Exp. Stat., South. Coop. Series Bull. No. 309, pp. 53-61

Mendenhall, W., Wackerly, D.D. and Scheaffer, R.L. 1990. Mathematical Statistics with Applications. PWS-Kent Publishing Company, Boston (Fourth Edition), 818p.

Murtha, J. and Willard, E. 1973. Statistics and Calculus a First Course. Prentice-Hall Inc., Englewood Cliffs, NJ, 529p.

Namkoong, G. 1979. Introduction to Quantitative Genetics. Technical Bulletin No. 1588, USDA Forest Service, 342p.

Patterson, H.D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. Biometrika 58:545-554

Phelps, W.R. and Czabator, F.L. 1978. Fusiform Rust of Southern Pines. Forest Insect and Disease Leaflet 26, USDA Forest Service, Washington, DC, 7p.

Pienaar, L.V., Page, H.H. and Rheney, J.W. 1990. Yield prediction for mechanically prepared slash pine plantations. Southern J. Appl. For. 14: 104-109

Pswarayi, I.Z. 1993. Genetic parameters and selection indices for a population of Pinus elliottii Engelm. var elliottii. Ph.D. Thesis, Oxford University, Linacre College, Oxford, 258p.

Riemenschneider, D.E. 1988. Heritability, age-age correlations, and inferences regarding juvenile selection in jack pine. For. Sci. 34:1076-1082

Rockwood, D.L. 1981. Volume Prediction for Genetically Improved Slash Pine Trees. Agric. Expt. Stat., Bulletin 819 (Technical), Inst. Food & Agric. Sci., Univ. Florida, Gainesville, 17p.

Rockwood, D.L. and Goddard, R.E. 1973. Predicted gains from fusiform rust resistance in slash pine. In: 12th South. For. Tree. Imp. Conf., June 12-13, Baton Rouge, LA., pp. 31-37.

SAS Institute Inc. 1988. SAS/STAT User's Guide, Release 6.03 Edition, SAS Institute Inc., Cary, NC, 1028p.

Samuel, C.J.A. 1991. The estimation of genetic parameters for growth and stem-form over 15 years in a diallel cross of sitka spruce. Silvae Genet. 40:67-72

Schaffer, H.E. and Usanis, R.A. 1969. General Least Squares Analysis of Diallel Experiments. A Computer Program - DIALL. Genetics Dept. Res. Rep. No.1, North Carolina State Univ., Raleigh NC, 61p.

Schmidt, R.A., Powers, H.R. and Snow, G.A. 1981. Application of genetic disease resistance for the control of fusiform rust in intensively managed southern pine. Phytopath. 71:993-997

Searle, S.R., Casella, G. and McCulloch, C.E. 1992. Variance Components. John Wiley & Sons, Inc., New York, 501p.

Shelbourne, C.J.A. 1972. Genotype-environment interaction: its study and implications in forest tree improvement. In: Proc. IUFRO Joint Genet. Symp. Jap. For. Expt. Station, Tokyo, Japan, pp. 1-28

Sluder, E.R. 1986. Gains from first-cycle selection in slash and longleaf pines. Silvae Genet. 35:155-159

Sluder, E.R. 1988. Inheritance and gain in a half-diallel cross among loblolly pines selected for resistance to fusiform rust. Silvae Genet. 37:22-26

Sluder, E.R. 1993. Results at age 15 years from a half-diallel cross among 10 loblolly pines selected for resistance to fusiform rust (Cronartium quercuum f. sp. fusiforme). Silvae Genet. 42:223-230

Snyder, E.B. and Namkoong, G. 1978. Inheritance in a diallel crossing experiment with longleaf pine. USDA For. Serv Res. Paper SO-140, South. For. Exp. Stn., New Orleans, LA 31p.

Sohn, S. and Goddard, R.E. 1979. Influence of infection percent on improvement of fusiform rust resistance in slash pine. Silvae Genet. 28:173-180

Sohn, S.I., Goddard, R.E. and Schmidt, R.A. 1975. Comparative performances of slash pine for fusiform rust resistance in high rust hazard locations. In: Proc. 13th South. For. Tree Imp. Conf., June 10-11, Raleigh, NC., pp. 204-211.

Sorensen, F.C. and White, T.L. 1988. Effect of natural inbreeding on variance structure in tests of wind-pollinated douglas-fir progenies. For. Sci. 34:102-118

Squillace, A.E. 1974. Average genetic correlations among offspring from open-pollinated forest trees. Silvae Genet. 23:149-156

Steel, G.D. and Torrie, J.H. 1981. Principles and Procedures of Statistics: A Biometrical Approach. McGraw-Hill Book Company, Singapore, (Second edition) 633p.

Swallow, W.H. and Monahan, J.F. 1984. Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. Technometrics 26:47-57

Swiger, L.A., Harvey, W.R., Everson, D.O. and Gregory, K.E. 1964. The variance of intraclass correlation involving groups with one observation. Biometrics 20:818-826

Talbert, C.B. 1984. An analysis of several approaches to multiple-trait index selection in loblolly pine (Pinus taeda L.). Ph.D. Thesis, Dept. For., North Carolina State Univ., Raleigh, NC, 106p.

Tallis, G.M. 1959. Sampling errors of genetic correlation coefficients calculated from analyses of variance and covariance. Aust. J. Stat. 1:35-43

Tauer, C.G. and McNew, R.W. 1985. Inheritance and correlation of growth of shortleaf pine in two environments. Silvae Genet. 34:5-11

Vargas-Hernandez, J. and Adams, W.T. 1991. Genetic variation of wood density components in young Douglas-fir: implications for tree breeding. Can. J. For. Res. 21: 1801-1807

Walkinshaw, C.H. and Anderson, R.L. 1988. How to classify fusiform rust galls on infected loblolly and slash pines. Protection Report R8-PR 14, USDA Forest Service, Southern Region, Atlanta, GA.

Walkinshaw, C.H. and Bey, C.F. 1981. Reaction of field-resistant slash pines to selected isolates of Cronartium quercuum f. sp. fusiforme. Phytopath. 71:1090-1092

Walkinshaw, C.H. and Roland, T.A. 1990. Incidence and histology of stem-girdling galls caused by fusiform rust. Phytopath. 80:251-255

Weisberg, S. 1985. Applied Linear Regression. John Wiley & Sons, Inc., New York, NY, (Second Edition), 324p.

Westfall, P.H. 1987. A comparison of variance component estimates of arbitrary underlying distributions. J. Amer. Stat. Assoc. 82:866-873

White, T.L. 1987. A conceptual framework for tree improvement programs. New Forests 4:325-342

White, T., Flinchum, M., Rockwood, D., Kok, H. and Powell, G. 1986. Twenty-Eighth Progress Report, Cooperative Forest Genetics Research Program, Dept. For., Univ. Florida, Gainesville, FL, 35p.

White, T.L. and Hodge, G.R. 1989. Predicting Breeding Values with Applications in Forest Tree Improvement. Forestry Sciences Series, Vol. 33, Kluwer Academic Publishers, Dordrecht, Netherlands, 367p.

White, T.L. and Hodge, G.R. 1992. Test designs and optimum age for parental selection in advanced generation progeny test of slash pine. Silvae Genet. 41(4-5):293-302

White, T.L, Hodge, G.R. and Powell, G.L. 1993. An advanced-generation tree improvement plan for slash pine in the southeastern United States. Silvae Genet. 42:359-371

White, T., Hodge, G., Rockwood, D., Kok, H., Powell, G., Flinchum, M., Sibley, J., Dippon, D. and Blakeslee, G. 1987. Twenty-Ninth Progress Report, Cooperative Forest Genetics Research Program, Dept. For., Univ. Florida, Gainesville, FL, 53p.

Winer, B.J., Brown, D.R. and Michels, K.M. 1991. Statistical Principles in Experimental Design. McGraw-Hill series in Psychology, McGraw-Hill Inc., New York, (Third Edition), 1057p.

Wright, A.J. 1985. Diallel designs, analyses, and reference populations. Heredity 54:307-311

Woolaston, R.R., Kanowski, P.J. and Nikles, D.G. 1990. Genetic parameter estimates for Pinus caribaea var. hondurensis in coastal Queensland, Australia. Silvae Genet. 39:21-28

Yamada, Y. 1962. Genotype by environment interaction and genetic correlation of the same traits under different environments. Japan J. Genetics 3:498-509

Yeh, F.C. and Heaman, C. 1982. Heritabilities and genetic and phenotypic correlations for height and diameter in coastal Douglas-fir. Can. J. For. Res. 12:181-185

BIOGRAPHICAL SKETCH

Mark John Dieters was born October, 1961, in Motueka, New Zealand. He and his family emigrated to Australia in October, 1961, and lived in and around Melbourne until December 1979, when he graduated from high school at St. Paul's College. He entered the Australian National University, Canberra, in January, 1980, to study forestry. In 1983 he completed his Bachelor of Science degree in forestry, and in 1984 commenced research for his honors degree with Dr. M.U. Slee on mangrove physiology at a field site near Darwin, Northern Territory. Mark was awarded a first class honors degree and won the Schlich medal in December, 1984, from the Department of Forestry, Australian National University. In 1985 he commenced work with the Queensland Forest Service, where he served as the officer in charge of the *Araucaria cunninghamii* breeding program until December, 1990. He entered the Doctor of Philosophy program at the University of Florida in January, 1991, and will return to Queensland Forest Service in 1994 where he will work with tropical conifers (including slash pine and its hybrids).

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Timothy L. White, Chairman
Professor of Forest Resources and Conservation

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Michael A. DeLorenzo
Associate Professor of Dairy Science

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gary R. Hodge
Associate Scientist of Forest Resources and Conservation

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ramon C. Littell
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Donald L. Rockwood
Professor of Forest Resources and Conservation

This dissertation was submitted to the Graduate Faculty of the School of Forest Resources and Conservation in the College of Agriculture and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1994

Loukas G. Arvanitis

Director, School of Forest Resources and Conservation

Dean, Graduate School